

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 1999, 11:20:12 ; Search time 13.2 Seconds  
(without alignments)  
16.150 Million cell updates/sec

Title: US-08-819-669D-26

Perfect score: 52

Sequence: 1 EADPTGHSY 9

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	9	1 R29769	Antigen E peptide.
2	52	100.0	9	1 R49224	HLA-A1 MAGE 1 anti
3	52	100.0	9	1 R47330	HLA-A1 MAGE 1 anti
4	52	100.0	9	1 R50281	MAGE-1 nonapeptide
5	52	100.0	9	1 R36675	Synthetic peptide
6	52	100.0	9	1 R5135	MAGE 1 immunogenic
7	52	100.0	309	1 R70909	Human melanoma ant
8	52	100.0	9	1 R65112	MAGE 1 immunogenic
9	52	100.0	9	1 R75954	Melanoma antigen (
10	52	100.0	9	1 R82988	P815 antigenic pep
11	52	100.0	9	1 R78824	MAGE-1 cytotoxic T
12	52	100.0	12	1 R80620	Immunogenic peptid
13	52	100.0	9	1 R83932	MHC class I restri
14	52	100.0	9	1 R90692	Human leukocyte an
15	52	100.0	9	1 R93343	MAGE-1 nonapeptide
16	52	100.0	9	1 W00897	Human melanoma MAG
17	52	100.0	10	1 W23038	MAGE-1/HLA-B44 tum
18	52	100.0	9	1 W36729	MAGE-1 antigenic p
19	52	100.0	9	1 W54622	Peptide from MAGE-
20	52	100.0	9	1 W68371	Human MAGE-1 pepti
21	52	100.0	9	1 W78838	MAGE-1 protein fra
22	52	100.0	9	1 W77125	gp75/TRP-1 synthet
23	52	100.0	9	1 W57536	Peptidase-resistan
24	52	100.0	9	1 W57536	Peptidase-resistan
25	52	100.0	309	1 W81548	Tumour rejection a
26	52	100.0	9	1 W08945	HLA-A1 binding pep
27	52	100.0	9	1 W00685	Tumour antigen boo
28	52	100.0	9	1 Y10633	Peptide antigen SE
29	52	100.0	9	1 Y10623	Peptide antigen SE
30	52	100.0	9	1 Y10424	HLA Class I motif
31	52	100.0	9	1 Y01727	Exemplary antigeni
32	49	94.2	9	1 R99342	HLA binding nonape
33	48	92.3	9	1 W57533	Peptidase-resistan
34	48	92.3	9	1 W57535	Peptidase-resistan
35	47	90.4	9	1 R99337	HLA binding nonape
36	47	90.4	9	1 R99339	HLA binding nonape
37	46	88.5	9	1 R99340	HLA binding nonape
38	44	84.6	9	1 R99338	HLA binding nonape
39	42	80.8	9	1 R99341	HLA binding nonape
40	42	80.8	9	1 Y10628	Peptide antigen SE
41	41.5	79.8	10	1 W54604	Peptide 1 from MAg
42	40	76.9	9	1 Y10629	Peptide antigen SE
43	37	71.2	88	1 W16332	Baboon MAGE-3 homo

#### ALIGNMENTS

```
RESULT 1
R29769
ID R29769 standard; Peptide; 9 AA.
AC R29769:
DT 22-APR-1993 (first entry)
DE Antigen E peptide.
KW Antigen; tumorigenic cell; A+ B+; T-cell; response; syngeneic;
KW animal; mouse; tumour rejection antigen precursor; TRAP; PIA.
OS Homo sapiens.
PN W09220356-A.
PD 26-NOV-1992.
PF 22-MAY-1992; 004354.
PR 23-MAY-1991; US-705702.
PR 09-JUL-1991; US-728838.
PR 23-SEP-1991; US-764364.
PR 12-DEC-1991; US-807043.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;
PI Van Den Eynde B, Van Der Bruggen P, Van Pel A;
DR WPI; 92-415460/50.
PT Nucleic acid mol. encoding a human tumour rejection antigen
PT precursor - useful as an immunostimulant in a vaccine for
PT treating and preventing cancers, also useful in diagnosis
PT Disclosure: Page 97; 142pp; English.
PS C: This sequence represents the sequence of the antigen E. Antigens such
as this one cause a T-cell response to be elicited which transplanted
CC into a syngeneic animal, usually a mouse. This antigen is derived from
CC the cell line MEL-3.1. See also Q33351.
CC Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
Db 1 EADPTGHSY 9

RESULT 2
R49224
ID R49224 standard; Protein; 9 AA.
AC R49224:
DT 31-AUG-1994 (first entry)
DE HLA-A1 MAGE 1 antigen peptide fragment 958.01.
KW Immunogenic; HLA-A3.2; HLA-A1; binding motif; MHC molecule;
KW immune response; viral infection; cancer; prostate cancer; lymphoma;
KW hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS Synthetic.
PN W09403205-A.
PD 17-FEB-1994.
PF 06-AUG-1993; 007421.
PR 07-AUG-1992; US-926666.
PR 05-MAR-1993; US-027746.
PA (CYTE-) CYTEL CORP.
PI Celis E, Grey HM, Kubo RT, Sette A;
PI WPI; 94-065403/08.
DR Peptide which specifically binds selected MHC allele - used to
PT induce an immune response for treatment or prevention of viral
PT infection or cancer, or for diagnosis.
PS Example 16; Page 116; 150pp; English.
CC The sequences given in R47304-33 and R49201-44 are immunogenic
CC peptides which have a HLA-A3.2, HLA-A1 or a HLA-A1 binding motif.
CC These peptides may be used in the composition of the invention.
CC These peptides are capable of binding selected MHC molecules and
```

CC inducing an immune response. They can be used to treat and/or  
 CC prevent viral infection and cancer, eg. prostate cancer, lymphoma,  
 CC hepatitis or AIDS. They can also be used to produce antibodies for  
 CC use as diagnostic or therapeutic agents. The peptides can also be  
 CC used as diagnostic agents.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 | | | | | | | | |  
 Db 1 EADPTGHSY 9

RESULT 3  
 R47330 ID R47330 standard; Protein; 9 AA.  
 AC R47330;  
 DE 31-AUG-1994 (first entry)  
 KW HLA-A1 MAGE 1 antigen peptide fragment 161-169.  
 KW immunogenic; HLA-A3.2; HLA-A1; binding motif; MHC molecule;  
 KW immune response; viral infection; cancer; prostate cancer; lymphoma;  
 KW hepatitis; AIDS; antibody; diagnosis; melanoma antigen.  
 OS Synthetic.  
 PN W09403205-A.  
 PD 17-FEB-1994.  
 PF 06-AUG-1993; U07421.  
 PR 07-AUG-1992; US-928666.  
 PR 05-MAR-1993; US-027746.  
 PA (CYTE-) CYTEL CORP.  
 PI Cellis E, Grey HM, Kubo RT, Sette A;  
 DR WPI; 94-065403/08.  
 PT Peptide which specifically binds selected MHC allele - used to  
 PT induce an immune response for treatment or prevention of viral  
 PT infection or cancer, or for diagnosis  
 PS Example 8; Page 52; 150pp; English.  
 CC The sequences given in R47304-33 and R49201-44 are immunogenic  
 CC peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.  
 CC These peptides may be used in the composition of the invention.  
 CC These peptides are capable of binding selected MHC molecules and  
 CC inducing an immune response. They can be used to treat and/or  
 CC prevent viral infection and cancer, eg. prostate cancer, lymphoma,  
 CC hepatitis or AIDS. They can also be used to produce antibodies for  
 CC use as diagnostic or therapeutic agents. The peptides can also be  
 CC used as diagnostic agents.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 | | | | | | | | |  
 Db 1 EADPTGHSY 9

RESULT 4  
 R50281 ID R50281 standard; Protein; 9 AA.  
 AC R50281;  
 DE 26-SEP-1994 (first entry)  
 KW MAGE-1 nonapeptide.  
 DE MAGE; nonapeptide; cancer; melanoma; breast cancer; HLA;  
 KW histocompatibility; human leucocyte antigen; probe; treatment;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN W09405304-A.  
 PD 17-MAR-1994.  
 PF 30-AUG-1993; U08157.

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 | | | | | | | | |  
 Db 1 EADPTGHSY 9

RESULT 5  
 R63675 ID R63675 standard; Protein; 9 AA.  
 AC R63675;  
 DE 22-JUN-1995 (first entry)  
 DE Synthetic peptide derived from exon 3.1 of MAGE 1.  
 KW Melanoma antigen-1; MAGE-1; cytolytic T cells; antigen E; exon 3.1.  
 OS Synthetic.  
 PN W09423031-A.  
 PD 13-OCT-1994.  
 PF 17-MAR-1993; U02877.  
 PR 26-MAR-1993; US-037230.  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PI Boon-falleur T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;  
 DR WPI; 94-333192/41.  
 PT New tumour rejection antigen precursor MAGE3 - useful in  
 PT treatment and diagnosis of cancer  
 PS Example 34; Page 36; 105pp; English.  
 CC R63675 is a synthetic peptide derived from exon 3.1 of melanoma  
 CC antigen-1 (MAGE-1). It was used to transfer antigen-E cytolytic T  
 CC lymphocyte sensitivity to normally non-sensitive cells.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 | | | | | | | | |  
 Db 1 EADPTGHSY 9

PR 31-AUG-1992; US-938334.  
 PR 26-MAR-1993; US-037230.  
 PR 07-JUN-1993; US-073103.  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PI Boon-falleur T, De Plaen E, Lurquin C, Traversari C;  
 PI Van Derbruggen P;  
 DR WPI; 94-100844/12.  
 DR N-PSDB; Q44751.  
 PT New nona-peptide derived from tumour rejection antigen precursor  
 PT - presented by HLA-A1 cancer cells, for use in diagnosis or  
 PT therapy of esp. melanoma and breast cancer.  
 PS Disclosure; Page 19; 33pp; English.  
 CC An isolated nonapeptide having the amino acid sequence Glu-Val-Asp-  
 CC Pro-Ile-Gly-His-Leu-Tyr is derived from the tumour rejection antigen  
 CC precursor encoded by the MAGE-3 gene and presented by HLA-A1. The  
 CC nonapeptide can be used in a vaccine to treat a cancerous condition  
 CC involving HLA-A1 subtype cancerous cells. The nucleic acid encoding  
 CC the nonapeptide can be used as a probe to identify tumour cells.  
 CC This sequence is homologous to the peptide described and is encoded  
 CC by the MAGE-1 gene.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 | | | | | | | | |  
 Db 1 EADPTGHSY 9

RESULT 6  
 R65135 ID R65135 standard; Peptide; 9 AA.  
 AC R65135;  
 DE 09-OCT-1995 (first entry)  
 DE MAGE 1 immunogenic peptide A01.  
 KW MAGE 1; immunogenic peptide A01; cytotoxic C cells;  
 KW in vitro activation; cancer; AIDS; bacterial infections; malaria;  
 OS Synthetic.  
 PN W09405304-A.  
 PD 17-MAR-1994.  
 PF 30-AUG-1993; U08157.

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 | | | | | | | | |  
 Db 1 EADPTGHSY 9

KW fungal infections; tuberculosis; hepatitis.  
 OS Homo sapiens.  
 PN WO9504817-A.  
 PD 16-FEB-1995.  
 PR 01-AUG-1994; U08672.  
 PF 06-AUG-1993; US-103401.  
 PA (CYTE-) CYTEL CORP.  
 PI Cellis E, Kubo R, Serra H, Tsai V, Wentworth P;  
 DR WPI: 95-090895/12.  
 PT In vitro activation of cytotoxic T cells for selected killing of  
 target cells - for treating e.g. cancer, AIDS, hepatitis etc. by  
 incubating them with antigen presenting cells loaded with  
 appropriate immunogenic peptide  
 PS Example 3; Page 38; 53pp; English.  
 CC R65109-R65145 are immunogenic peptides, they are used in a new  
 method for the in vitro activation of cytotoxic T cells (CTC).  
 CC This is achieved by incubating the CTCs with antigen presenting  
 cells loaded with an appropriate immunogenic peptide (e.g. one  
 of the above peptides). By selecting the peptides used the  
 following diseases and infections can be treated; cancer, AIDS,  
 hepatitis, other viral and bacterial infections, malaria and  
 tuberculosis.  
 CC Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 |||||  
 DB 1 EADPTGHSY 9

RESULT 7  
 R70909 ID R70909 standard; Protein: 309 AA.  
 AC R70909;  
 DT 09-OCT-1995 (first entry)  
 DE Human melanoma antigen MAGE-1.  
 KW Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;  
 KW HLA-restricted cytotoxic T-lymphocyte activity.  
 OS Homo sapiens.  
 PN WO9504542-A.  
 PD 16-FEB-1995.  
 PR 02-AUG-1994; U08721.  
 PF 06-AUG-1993; US-103623.  
 PA (CYTE-) CYTEL CORP.  
 PI Fikes JD, Livingston BD, Sette AD, Sidney JC;  
 DR WPI: 95-090681/12.  
 DR N-PSDB; Q85435.  
 PT Human melanoma antigen, MAGE-1, peptide(s) - useful for  
 stimulating immune response against melanoma  
 PS Example 1; Fig 1; 59pp; English.  
 CC Q85435 encodes R70909 human melanoma antigen MAGE-1, it was used  
 to produce the C-terminal MAGE-1 peptides described in R70915 to  
 R70969. These peptides are useful for defining epitopes that  
 engender a HLA-restricted cytotoxic lymphocyte activity against  
 MAGE-1 antigens. Compsns. containing these peptides can be  
 administered, as a vaccine to patients susceptible to MAGE  
 associated tumours, e.g. melanomas.  
 CC Sequence 309 AA;

Query Match 100.0%; Score 52; DB 1; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 0.016;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 |||||  
 DB 161 EADPTGHSY 169

RESULT 8  
 R65112 ID R65112 standard; peptide: 9 AA.  
 AC R65112;  
 DT 06-OCT-1995 (first entry)  
 DE MAGE 1 immunogenic peptide 161-169  
 KW MAGE 1; immunogenic peptide 161-169; cytotoxic C cells;  
 KW in vitro activation; cancer; AIDS; bacterial infections; malaria;  
 KW fungal infections; tuberculosis; hepatitis.  
 OS Homo sapiens.  
 PN WO9504817-A.  
 PD 16-FEB-1995.  
 PR 01-AUG-1994; U08672.  
 PF 06-AUG-1993; US-103401.  
 PA (CYTE-) CYTEL CORP.  
 PI Cellis E, Kubo R, Serra H, Tsai V, Wentworth P;  
 DR WPI: 95-090895/12.  
 PT In vitro activation of cytotoxic T cells for selected killing of  
 target cells - for treating e.g. cancer, AIDS, hepatitis etc. by  
 incubating them with antigen presenting cells loaded with  
 appropriate immunogenic peptide  
 PS Example 3; Page 35; 53pp; English.  
 CC R65109-R65145 are immunogenic peptides, they are used in a new  
 method for the in vitro activation of cytotoxic T cells (CTC).  
 CC This is achieved by incubating the CTCs with antigen presenting  
 cells loaded with an appropriate immunogenic peptide (e.g. one  
 of the above peptides). By selecting the peptides used the  
 following diseases and infections can be treated; cancer, AIDS,  
 hepatitis, other viral and bacterial infections, malaria and  
 tuberculosis.  
 CC Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 |||||  
 DB 1 EADPTGHSY 9

RESULT 9  
 R73954 ID R73954 standard; Peptide: 9 AA.  
 AC R73954;  
 DT 06-MAR-1996 (first entry)  
 DE Melanoma antigen (MAGE-1) epitope.  
 KW MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;  
 KW cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.  
 OS Homo sapiens.  
 PN WO9519783-A1.  
 PD 27-JUL-1995.  
 PF 25-JAN-1995; U01000.  
 PR 25-JAN-1994; US-186266.  
 PA (CYTE-) CYTEL CORP.  
 PI Cellis E, Grey HM, Kubo RT, Sette A;  
 DR WPI: 95-269270/35.  
 PT Immunogenic peptide(s) that induce immune response to cancer cells  
 - that express a MAGE-3 protein peptide epitope used in vaccines or  
 adoptive immunotherapy to induce cytotoxic T lymphocytes  
 PS Example; Page 33; 44pp; English.  
 CC R73954 is derived from MAGE-1 protein. It was used to show the  
 specificity of CTL response to MAGE-3 peptides shown in R75942-53.  
 CC R75942 is derived from the sequence of the melanoma antigen (MAGE-3)  
 protein and can be used to elicit a primary cytotoxic T lymphocyte  
 response against cells expressing MAGE-3. Synthetic peptides R75945-53  
 can be used therapeutically to elicit CTL responses to melanoma, breast,  
 colon, prostate, or other cells which express proteins with this epitope.  
 CC The peptides have specific HLA-A1 binding capacity.  
 CC Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 |||||  
 DB 1 EADPTGHSY 9

RESULT 10  
 R82988  
 ID R82988 standard; Peptide; 9 AA.  
 AC R82988;  
 DT 26-FEB-1996 (first entry)  
 DE P815 antigenic peptide.  
 KW P815 antigen; P1A antigen; cancer; vaccine.  
 OS Synthetic.  
 PN WO9523874-A1.  
 PD 08-SEP-1995.  
 PF 23-FEB-1995; U02203.  
 PR 01-MAR-1994; US-204727.  
 PR 10-MAR-1994; US-209172.  
 PR 01-SEP-1994; US-299849.  
 PR 30-NOV-1994; US-346774.  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PI Boon-Falleur T, Brasseur F, Chomez P, De Plaen E;  
 PI De Smet C, Gaugler B, Lethe B, Marchand M, Patard J;  
 PI Szikora J, Van Den Eynde B, Van Derbruggen P, Weynants P;  
 DR WPI: 95-320586/41.  
 PT Determ. of cancerous condition(s) - using a nucleic acid as a  
 PT primer to determine expression of a MAGE tumour rejection antigen  
 PT precursor  
 PS Example 13; Page 22; 121pp; English.  
 CC Using the sequence of the P815A antigen precursor gene P1A  
 CC (T0117), an antigenic peptide (R82988) which was A+B+ (i.e.  
 CC characteristic of cells which express both A and B antigens) was  
 CC produced. The peptide lysed PO.HTR cells in the presence of  
 CC cytolytic T lymphocyte cell lines, and may be useful as a vaccine  
 CC component.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 |||||  
 DB 1 EADPTGHSY 9

RESULT 11  
 R78824  
 ID R78824 standard; peptide; 9 AA.  
 AC R78824;  
 DT 26-MAR-1996 (first entry)  
 DE MAGE-1 cytotoxic T lymphocyte epitope.  
 KW MAGE-1; cytotoxic T; CTL; epitope; helper T; HTL; lymphocyte;  
 KW cell; viruses; parasites; tumours; antigens; disease prevention;  
 KW treatment.  
 OS Homo sapiens.  
 PN WO9522317-A1.  
 PD 24-AUG-1995.  
 PF 16-FEB-1995; U02121.  
 PR 16-FEB-1994; US-197484.  
 PA (CITE-) CITE CORP.  
 PI Cells E, Chesnut RW, Grey H, Sette AD, Vitiello MA;  
 DR WPI: 95-302545/39.  
 PT Compn. inducing cytotoxic T lymphocyte response to pref. viral,  
 PT bacterial, parasitic or tumour antigens - useful in the treatment  
 PT and prevention of diseases associated with the antigen e.g.  
 PT hepatitis B  
 PS Disclosure; Page 17; 109pp; English.

CC A compsn. which induces a cytotoxic T lymphocyte (CTL) response to  
 CC an antigen (Ag) in a mammal comprises, a CTL Ag response inducing  
 CC peptide (i.e. R78824-R78853) and a lipid conjugated helper T cell  
 CC inducing peptide. The compsn. induces a CTL response to bacterial,  
 CC viral or tumour Ags, and is therefore useful in the treatment and  
 CC prevention of diseases associated with the Ag.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 |||||  
 DB 1 EADPTGHSY 9

RESULT 12  
 R80620  
 ID R80620 standard; Protein; 12 AA.  
 AC R80620;  
 DT 28-FEB-1996 (first entry)  
 DE Immunogenic peptide of tumour rejection antigen (MAGE-1).  
 KW Tumour rejection; antigen; MAGE-1; monoclonal antibody; MAb;  
 KW diagnosis; immunoassay; cancer; immunogen; antisera.  
 OS Homo sapiens.  
 PN WO9520974-A1.  
 PD 10-AUG-1995.  
 PF 05-JAN-1995; U00095.  
 PR 01-FEB-1994; US-190411.  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PA (SLOK) SLOAN KETTERING INST CANCER RES.  
 PA (SLOK) MEMORIAL-SLOAN-KETTERING CANCER CENT.  
 PI Boon-falleur T, Chen Y, Garin-Chesa P, Old LJ, Rettig WJ;  
 DR WPI: 95-283606/37.  
 PT New monoclonal antibody binding specifically to MAGE-1 - useful for  
 PT diagnosis and monitoring of cancer, also new hybridomas, recombinant  
 PT MAGE-1 and immunogenic peptide(s)  
 PS Claim 12; Page 20; 33pp; English.  
 CC A monoclonal antibody directed against the tumour rejection antigen  
 CC (MAGE-1) can be used to detect MAGE-1 in samples by standard  
 CC immunoassay methods for diagnosis and monitoring of cancer etc. The  
 CC monoclonal antibody is designated MA454 and is produced by the  
 CC hybridoma deposited as ATCC HB11540. The monoclonal antibody is  
 CC specific for MAGE-1, having no reactivity for MAGE-2 or MAGE-3.  
 CC Peptide fragments of MAGE-1 (see R80618-20) may be useful as  
 CC immunogens for production of the monoclonal antibody and antisera.  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00047;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
 |||||  
 DB 4 EADPTGHSY 12

RESULT 13  
 R83932  
 ID R83932 standard; peptide; 9 AA.  
 AC R83932;  
 DT 05-JUN-1996 (first entry)  
 DE MHC class I restricted antigenic peptide #2.  
 KW MHC class I; antigen; MAGE; melanoma; breast cancer; bladder cancer;  
 KW Titermax; cytotoxic T-lymphocyte; tumour; pathogenic disease; bacteria;  
 KW parasite; human; animal.  
 OS Synthetic.  
 PN WO9528958-A1.  
 PD 02-NOV-1995.



```
PF 21-APR-1995; U04975.
PR 22-APR-1994; US-233496.
PA (SLOK ) SLOAN KETTERING INST CANCER RES.
PI DYALL R, NIKOLIC-ZUGIC J;
DR WPI: 95-382848/49.
PT Cytotoxic T-cell induction by MHC class I-restricted peptide in
PT adjuvant - useful for treating tumours and bacterial or parasitic
PT pathogenic diseases
PS Claim 11; Page 38; 50pp; English.
CC The sequences given in R83931-49 are MHC class I restricted 8-12
CC amino acid antigenic peptides. This peptide is derived from MAGE
CC and is present in melanoma, breast and bladder cancer. These
CC peptides may be administered to a subject in combination with a
CC suitable adjuvant, pref. Titermax (RTM), to induce cytotoxic T-
CC lymphocytes. This method may be used in the treatment of a tumour
CC or a pathogenic disease, esp. diseases of bacterial or parasitic
CC origin, in humans and animals, e.g. monkeys, dogs, cows, horses, etc.
CC Sequence 9 AA;
SQ

Query Match 100.0%; Score 52; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
Db 1 EADPTGHSY 9

RESULT 14
R90692
ID R90692 standard; peptide; 9 AA.
AC R90692;
DT 31-JUL-1996 (first entry)
DE Human leukocyte antigen (HLA-A1) presented peptide M23-E.
KW Human leukocyte antigen; HLA-A1; MAGE-1 derived;
KW blood mononuclear cell; BMC; CD8-beta+ cell; cytolytic T cell;
KW CTL cell; treatment; tumour cell; diagnosis; assay;
KW presented peptide.
OS Synthetic.
PN W09535500-A1.
PD 28-DEC-1995.
PF 14-JUN-1995; U07559.
PR 17-JUN-1994; US-261541.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-Palleur T, Coulie P, Van Der Bruggen P;
DR WPI: 96-058510/06.
PT Prodn. of specific cytolytic T cell sub-populations - by contacting
PT blood mononuclear cells with specific peptide(s) and a population of
PT CD8-beta(+) cells
PS Claim 5; Page 19; 25pp; English.
CC The present peptide is the human leukocyte antigen (HLA-A1), MAGE-1
CC derived presented peptide, M23-E. By contacting a sample of blood
CC mononuclear cells (BMC) with the peptide (which binds directly to
CC HLA-A1 mols. on the surface of the BMC) and CD8-beta+ cells (which
CC stimulate peptide/HLA-A1 complex specific CD8-beta+ cells), a
CC peptide/HLA-A1 complex specific cytolytic T (CTL) cell
CC subpopulation can be obt'd. The CTL cells obt'd. can be
CC administered to a patient to treat tumour cell related conditions,
CC and can be used in diagnostic methods, e.g. in assays for the
CC peptide/HLA-A1 complex.
SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
Db 1 EADPTGHSY 9

RESULT 15
R99343
ID R99343 standard; Protein; 9 AA.
AC R99343;
DT 22-APR-1997 (first entry)
DE MAGE-1 nonapeptide.
KW HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW therapy.
OS Homo sapiens.
PN W09626214-A1.
PD 29-AUG-1996.
PF 01-FEB-1995; U01489.
PR (LUDW-) LUDWIG INST CANCER RES.
PA Boon-Palleur T, De Plaen E, Gaugler B, Lurquin C;
PI Boonero P, Traversari C, Van Den Eynde B, Van Der Bruggen P;
DR WPI: 96-402317/40.
DR N-PSDB; T35408.
PT New nona-peptide(s) that bind to HLA molecule(s) and induce lysis
PT by specific cytolytic T cells, for diagnosis and treatment of
PT tumours and to expand T cells in vitro.
PS Example 4; Fig 4; 41pp; English.
CC R99343-R99350 represent MAGE nonapeptides, based on the tumour rejection
CC antigen region of the full length MAGE sequences. These peptides were
CC used to design the nonapeptides of the invention (see R99337-R99342),
CC which bind to a HLA molecule on a cell, and provoke lysis by cytolytic T
CC cells (CTLs) specific for a complex of the HLA molecule and nonapeptide.
CC The nonapeptides can be used diagnostically to identify tumours
CC expressing a particular HLA molecule, or to identify cancer cells. The
CC peptides can also be used therapeutically, to induce a CTL response to
CC tumours (where the peptides are optionally coupled to tumour-specific
CC antibodies), or to induce a response by CTLs that are otherwise inactive.
CC The peptide sequences may also be used to expand specific CTLs in vitro
CC for later return to the patient, such as for treating melanoma. Tumour
CC cells can be identified by using DNA encoding the nonapeptides as probes.
CC Non-human cells transformed with the HLA-A1 gene and a DNA sequence
CC encoding one of the peptides, can be used to generate CTLs, or to detect
CC the presence of CTLs in human samples. The non-human transformed cells,
CC when polytransformed, are universal effector cells, and can be used in
CC vaccines, or for treating melanoma or breast cancer.
SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
Db 1 EADPTGHSY 9

Search completed: December 10, 1999, 14:27:26
Job time: 11234 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 1999, 07:27:34 ; Search time 10.2 Seconds  
(without alignments)  
35.352 Million cell updates/sec

Title: US-08-819-669D-26  
Perfect score: 52  
Sequence: 1 EADPTGHSY 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_50.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	52	100.0	280	2 JC2358
2	44	84.6	369	2 138659
3	43	82.7	315	2 138668
4	43	82.7	234	2 138667
5	42	80.8	319	2 138660
6	37	71.2	555	1 RGASWA
7	36	69.2	3396	1 A42551
8	36	69.2	314	2 JC2361
9	36	69.2	314	2 JC2360
10	36	69.2	925	2 PH1299
11	35	67.3	925	2 A39216
12	35	67.3	1033	2 S02168
13	35	67.3	98	2 F70769
14	34	65.4	497	2 S33938
15	34	65.4	370	2 S49008
16	34	65.4	1670	2 S71551
17	34	65.4	197	2 A70832
18	34	65.4	878	2 S44543
19	33	63.5	488	1 S55874
20	33	63.5	488	1 A53107
21	33	63.5	372	2 S32581
22	33	63.5	775	2 S57920
23	33	63.5	2123	2 S55089
24	33	63.5	295	2 C69180
25	33	63.5	288	2 A56279
26	33	63.5	747	2 B47093
27	33	63.5	597	2 S37849
28	33	63.5	1040	2 A38306
29	33	63.5	301	2 C71194
30	32	61.5	133	1 F45345
31	32	61.5	640	1 2VBPT5
32	32	61.5	283	2 E59626
33	32	61.5	533	2 S71778
34	32	61.5	669	2 I38029
35	32	61.5	349	2 A28658
36	32	61.5	506	2 S62629
37	32	61.5	381	2 A37276
38	32	61.5	379	2 A42421
39	32	61.5	700	2 S38928

multidrug resistant  
profilaggrin - hum  
tumor-associated a  
conserved hypotet  
toxin-like outer m  
tyrosine kinase su

## ALIGNMENTS

## RESULT 1

JC2358

tumor-associated antigen MAGE-1 - human

C:Species: Homo sapiens (man)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 04-Sep-1998

C:Accession: JC2358

R: Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.

Biochem. Biophys. Res. Commun. 202, 549-555, 1994

A:Title: Cloning and analysis of MAGE-1-related genes.

A:Reference number: JC2358; MUID:94311935

A:Accession: JC2358

A:Molecule type: mRNA

A:Residues: 1-280 <DIN>

A:Experimental source: melanoma cell line DML50

C:Genetics:

A:Gene: MAGE

C:Superfamily: tumor associated protein MAGE

F:161-169/Region: HLA-A1 binding #status predicted

Query Match 100.0%; Score 52; DB 2; Length 280;  
Best Local Similarity 100.0%; Pred. No. 0.0099;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9

Db 161 EADPTGHSY 169

## RESULT 2

I38659

MAGE-10 antigen - human

C:Species: Homo sapiens (man)

C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 04-Sep-1998

C:Accession: I38659

R: De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.;

Immunogenetics 40, 360-369, 1994

A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE

A:Reference number: I38659; MUID:95012457

A:Accession: I38659

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-369 <RES>

A:Cross-references: EMBL:U10685; NID:9533510; PID:9533511

C:Genetics:

A:Gene: GDB:MAGEA10; MAGE10

A:Cross-references: GDB:331126

A:Map position: Xq28-Xq28

A:Introns: #status absent

C:Superfamily: tumor associated protein MAGE

Query Match 84.6%; Score 44; DB 2; Length 369;  
Best Local Similarity 77.8%; Pred. No. 0.45;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9

Db 193 EVDPTGHSF 201

RESULT 3  
I38668  
MAGE-9 antigen - human  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 04-Sep-1998  
C:Accession: I38668  
R:De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; Br  
con, T.  
Immunogenetics 40, 360-369, 1994  
A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE fam  
A:Reference number: I38659; MUID:95012457  
A:Accession: I38668  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-315 <RES>  
A:Cross-references: EMBL:U10694; NID:g533527; PID:g533528  
C:Genetics:  
A:Gene: GDB:MAGEA9; MAGE9  
A:Cross-references: GDB:331125  
A:Map position: Xp21.3-Xp21.3  
A:Introns: #status absent  
C:Superfamily: tumor associated protein MAGE

Query Match 82.7%; Score 43; DB 2; Length 315;  
Best Local Similarity 77.8%; Pred. No. 0.6;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
| | | | |  
DB 167 EVDPAHSHY 175

RESULT 4  
I38667  
MAGE-8 antigen - human  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 04-Sep-1998  
C:Accession: I38667  
R:De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; Br  
con, T.  
Immunogenetics 40, 360-369, 1994  
A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE fam  
A:Reference number: I38659; MUID:95012457  
A:Accession: I38667  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-234 <RES>  
A:Cross-references: EMBL:U10693; NID:g533525; PID:g533526  
C:Genetics:  
A:Gene: GDB:MAGEA8; MAGE8  
A:Cross-references: GDB:331123  
A:Map position: Xq28-Xq28  
A:Introns: #status absent  
C:Superfamily: tumor associated protein MAGE

Query Match 82.7%; Score 43; DB 2; Length 234;  
Best Local Similarity 77.8%; Pred. No. 0.44;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1-EADPTGHSY 9  
| | | | |  
DB 171 EVDPAHSHY 179

RESULT 5  
I38660  
MAGE-11 antigen - human  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 04-Sep-1998  
C:Accession: I38660  
R:De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; Br

con, T.  
Immunogenetics 40, 360-369, 1994  
A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE  
A:Reference number: I38659; MUID:95012457  
A:Accession: I38660  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-319 <RES>  
A:Cross-references: EMBL:U10686; NID:g533512; PID:g533513  
C:Genetics:  
A:Gene: GDB:MAGE11; MAGE11  
A:Cross-references: GDB:331128  
A:Map position: Xq28-Xq28  
A:Introns: #status absent  
C:Superfamily: tumor associated protein MAGE

Query Match 80.8%; Score 42; DB 2; Length 319;  
Best Local Similarity 77.8%; Pred. No. 0.94;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9  
| | | | |  
DB 171 EVDPTSHSY 179

RESULT 6  
RGASWA  
regulatory protein weta - Emericella nidulans  
C:Species: Emericella nidulans, Aspergillus nidulans  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 24-Oct-1997  
C:Accession: A39665  
R:Marshall, M.A.; Timberlake, W.E.  
Mol. Cell. Biol. 11, 55-62, 1991  
A:Title: Aspergillus nidulans weta activates spore-specific gene expression.  
A:Reference number: A39665; MUID:91094871  
A:Accession: A39665  
A:Molecule type: DNA  
A:Residues: 1-555 <MAR>  
A:Cross-references: GB:M60528; GB:M35758; NID:g168108; PID:g168109  
C:Comment: The products of the genes brIA, abaa, and weta are required for activation  
C:Genetics:  
A:Gene: weta  
C:Superfamily: regulatory protein weta  
C:Keywords: transcription regulation

Query Match 71.2%; Score 37; DB 1; Length 555;  
Best Local Similarity 87.5%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHS 8  
| | | | |  
DB 109 EADATGHS 116

RESULT 7  
A42551  
genome polyprotein - dengue virus type 1 (strain Singapore S275/90)  
N:Contains: capsid protein; envelope protein; membrane protein; nonstructural protein  
a: nonstructural protein NS4b; nonstructural protein NS5  
C:Species: dengue virus type 1  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 15-May-1998  
C:Accession: A42551  
R:Fu, J.; Tan, B.H.; Yap, E.H.; Chan, Y.C.; Tan, Y.H.  
Virology 188, 953-958, 1992  
A:Title: Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90).  
A:Reference number: A42551; MUID:92263809  
A:Accession: A42551  
A:Molecule type: genomic RNA  
A:Residues: 1-3396 <FUJ>  
A:Cross-references: GB:M87512  
C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology

C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
 F:1-114/Product: capsid protein #status predicted <CAP>  
 F:115-281/Product: membrane protein precursor #status predicted <MEP>  
 F:115-204/Domain: nonterminal signal sequence #status predicted <SIG>  
 F:205-281/Product: membrane protein #status predicted <MEM>  
 F:267-279/Domain: transmembrane #status predicted <TM1>  
 F:282-774/Product: envelope protein #status predicted <ENV>  
 F:753-769/Domain: transmembrane #status predicted <TM2>  
 F:775-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
 F:1128-1344/Product: nonstructural protein NS2a #status predicted <NS2A>  
 F:1345-1474/Product: nonstructural protein NS2b #status predicted <NS2B>  
 F:1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>  
 F:1668-1938/Domain: DEAD/H box helicase homology <DEAD>  
 F:1668-1675/Region: nucleotide-binding motif A (P-loop)  
 F:1759-1760/Region: nucleotide-binding motif B  
 F:1759-1762/Region: DEAD motif  
 F:2094-2243/Product: nonstructural protein NS4a #status predicted <NS4A>  
 F:2244-2492/Product: nonstructural protein NS4b #status predicted <NS4B>  
 F:2493-3396/Product: nonstructural protein NS5 #status predicted <NS5>  
 F:183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.2%; Score 36; DB 1; Length 3396;  
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHS 8

|||||

Db 3383 ESDPKGHS 3390

RESULT 8

tumor-associated antigen MAGE-3 - human  
 N:Alternate names: MAGE 3 protein  
 C:Species: Homo sapiens (man)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 04-Sep-1998  
 C:Accession: JC2361; PH1296; 138438  
 R:Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.  
 Biochem. Biophys. Res. Commun. 202, 549-555, 1994  
 A:Title: Cloning and analysis of MAGE-1-related genes.  
 A:Reference number: JC2358; MUID:93018875  
 A:Accession: JC2361  
 A:Molecule type: mRNA  
 A:Residues: 1-314 <DIN>  
 A:Experimental source: melanoma cell line DM150  
 R:Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van P.  
 J. Exp. Med. 176, 1453-1457, 1992  
 A:Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic  
 A:Reference number: PH1294; MUID:93018875  
 A:Accession: PH1296  
 A:Molecule type: DNA  
 A:Residues: 168-176 <TRA>  
 R:Gaigler, B.; Van den Eynde, B.; van der Bruggen, P.; Romero, P.; Gaforio, J.J.; De Pla  
 J. Exp. Med. 179, 921-930, 1994  
 A:Title: Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous c  
 A:Reference number: I38438; MUID:94157413  
 A:Accession: I38438  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-314 <RES>  
 A:Cross-references: EMBL:U03735; NID:9468825; PID:9468826  
 C:Genetics:  
 A:Gene: MAGE-3  
 C:Superfamily: tumor associated protein MAGE  
 F:168-176/Region: HLA-A1 binding #status predicted

Query Match 69.2%; Score 36; DB 2; Length 314;  
 Best Local Similarity 66.7%; Pred. No. 13;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHS 9

Db 168 EVDPIGHLY 176

RESULT 9

tumor-associated antigen, MAGE 6 - human  
 N:Alternate names: melanoma antigen 6; tumor-associated antigen, MAGE-3b  
 C:Species: Homo sapiens (man)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 04-Sep-1998  
 C:Accession: JC2360; PH1301; I38665; G01445  
 R:Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.  
 Biochem. Biophys. Res. Commun. 202, 549-555, 1994  
 A:Title: Cloning and analysis of MAGE-1-related genes.  
 A:Reference number: JC2358; MUID:93018875  
 A:Accession: JC2360  
 A:Molecule type: mRNA  
 A:Residues: 1-314 <DIN>  
 A:Experimental source: melanoma cell line DM150  
 R:Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van P.  
 J. Exp. Med. 176, 1453-1457, 1992  
 A:Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic  
 A:Reference number: PH1294; MUID:93018875  
 A:Accession: PH1301  
 A:Molecule type: DNA  
 A:Residues: 168-176 <TRA>  
 R:De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.;  
 Immunogenetics 40, 360-369, 1994  
 A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE  
 A:Reference number: I38659; MUID:95012457  
 A:Accession: I38665  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-314 <RES>  
 A:Cross-references: EMBL:U10691; NID:9533522; PID:9533523  
 R:Fenton, R.G.  
 submitted to the EMBL Data Library, June 1994  
 A:Reference number: G07126  
 A:Accession: G01445  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-314 <FEN>  
 A:Cross-references: EMBL:U10339; NID:9499121; PID:9499122  
 C:Genetics:  
 A:Gene: GDB:MAGEA6; MAGE6  
 A:Cross-references: GDB:331121  
 Z:Map position: Xq28-Xq28  
 A:Introns: #status absent  
 C:Superfamily: tumor associated protein MAGE  
 F:168-176/Region: HLA-A1 binding #status predicted

Query Match 69.2%; Score 36; DB 2; Length 314;  
 Best Local Similarity 66.7%; Pred. No. 13;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHS 9

|||||

Db 168 EVDPIGHVY 176

RESULT 10

MAGE 5 protein - human (fragment)  
 N:Alternate names: MAGE 51 protein  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 03-Aug-1998  
 C:Accession: PH1299; PH1300  
 R:Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van P.  
 J. Exp. Med. 176, 1453-1457, 1992  
 A:Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic  
 A:Reference number: PH1294; MUID:93018875

A;Accession: PH1299  
A;Molecule type: DNA  
A;Residues: 1-9 <TRA>  
A;Accession: PH1300  
A;Molecule type: DNA  
A;Residues: 1-9 <TR2>

Query Match 69.2%; Score 36; DB 2; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.2e+05;  
Matches 6; Conservative 2; Mismatches 1; Indels

```
QY      1 EADPTGHSY 9
        ||||| : : |
Dp      1 EADPTSNY 9
```

RESULT 11  
A39216  
plasma cell membrane glycoprotein PC-1 - human  
N:Contains: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.4.1)  
C:Species: Homo sapiens (man)  
C:Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 29-Aug-1997  
C:Accession: A39216; S21706; S23587; S51030  
R:Buckley, M.F.; Loveland, K.A.; McKinstry, W.J.; Garson, O.M.; Goding, J.W.  
J. Biol. Chem. 265, 17506-17511, 1990  
A:Title: Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human molecule, and  
A:Reference number: A39216; MUID:91009202  
A:Accession: A39216  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-925 <BUC>  
A:Cross-references: GB:J05654  
R:Funakoshi, I.; Kato, H.; Horie, K.; Yano, T.; Hori, Y.; Kobayashi, H.; Inoue, T.; Suzuki, T.  
Arch. Biochem. Biophys. 295, 180-187, 1992  
A:Title: Molecular cloning of cDNAs for human fibroblast nucleotide pyrophosphatase.  
A:Reference number: S21706; MUID:92246539  
A:Accession: S21706  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-925 <FUNI>  
A:Accession: S23587  
A:Molecule type: protein  
A:Residues: 116-121;247-271, 'X', 273-275;279-280, 'X', 282-283;303-316;362-364;449-465;482-483  
A>Note: it is uncertain whether Met-1 or Met-53 is the initiator

```

Query Match      67.3%; Score 35; DB 2; Length 925;
Best Local Similarity 66.7%; Pred. No. 63;
Matches         6: Conservative 1; Mismatches 2; Indels

```

Qy 1 EADPTGHSY 9  
| | : | | |

Db 374 EPDSSGHSY, 382.

RESULT  
S02168

type I site-specific deoxyribonuclease (EC 3.1.21.3) EcoRI24/3 chain hsdR - Escherichia coli  
N:Alternate names: type I restriction enzyme EcoRI24/3 chain hsdR  
C:Species: Escherichia coli  
C:Date: 01-dec-1989 #sequences\_revision 01-Dec-1989 #text\_change 09-Sep-1997  
C:Accession: S02168  
R:Price, C.; Lingner, J.; Bickle, T.A.  
J. Mol. Biol. 205, 115-125, 1989  
A:Title: Basis for changes in DNA recognition by the EcoRI24 and EcoRI24/3 type I DNA  
A:Reference number: S02165; MUID: 89178628

Query Match	67.3%	Score 35;	DB 2;	Length 1033;
Best Local Similarity	56.7%	Pred. No. 70;		
Matches	6;	Conservative	2;	Mismatches
				1; Indels
				0; Gaps
				0;

Qy	1	EADPTGHSY	9.
		: :	
Db	22	KAEP TGDSY	30

RESULT 13  
F70769

hypothetical protein Rv1322 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 17-Jul-1998  
C:Accession: F70769  
R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Title: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Author: Deciphering the Biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MUID:98295987  
A:Accession: F70769  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-98 <COL>  
A:Cross-references: GB: Z73902; GB: AL123456; NID: G3261576; PID: e245016; PID: g1340088  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: Rv1322

Query Match . 67.3%; Score 35; DB 2; Length 98;  
Best Local Similarity 66.7%; Pred. No. 6.1;  
Matches 6: Conservative 0; Mismatches 3: Indels

Qy	1	EADPTGHSY	9
Db	24	EAGPDGHEY	32

RESULT	14
S33938	

penton protein (III) - human adenovirus 12  
C:Species: Mastadenovirus h12 (human adenovirus 12)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Sep-1997  
C:Accession: S33938  
R:Sprengel, J.  
submitted to the EMBL Data Library. June 1993

A:Reference number: S33928  
A:Accession: S33938  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-497 <SPR>  
A:Cross-references: EMBL:X73487; NID:g313361; PID:g313372  
C:Superfamily: adenovirus penton protein

Query Match 65.4%; Score 34; DB 2; Length 497;  
Best Local Similarity 66.7%; Pred. No. 51;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9  
| | | | |  
Db 310 ETPKGRSY 318

RESULT 15  
S49008  
fork head protein - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 17-Mar-1999  
C:Accession: S49008; B56556  
R:Lef, J.; Clement, J.H.; Oschwald, R.; Koester, M.; Knoechel, W.  
Mech. Dev. 45, 117-126, 1994  
A:Title: Spatial and temporal transcription patterns of the forkhead related XFD-2/XFD-2  
A:Reference number: S49008; MUID:94257528  
A:Accession: S49008  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-370 <LEF>  
A:Cross-references: EMBL:X74315; NID:g511159; PID:g511160  
R:Knoechel, S.; Lef, J.; Clement, J.; Klocke, B.; Hille, S.; Koster, M.; Knoechel, W.  
Mech. Dev. 38, 157-165, 1992  
A:Title: Activin A induced expression of a fork head related gene in posterior chordames  
A:Reference number: A56556; MUID:93041288  
A:Accession: B56556  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 118-228 <KNO>  
A:Experimental source: gastrula  
A:Note: sequence extracted from NCBI backbone (NCBIP:118178)  
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology  
F:127-218/Domain: fork head DNA-binding domain homology <FHD>

Query Match 65.4%; Score 34; DB 2; Length 370;  
Best Local Similarity 75.0%; Pred. No. 38;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHS 8  
| | | | |  
Db 257 ELSPTGHS 264

Search completed: December 11, 1999, 07:55:08  
Job time: 1654 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 1999, 19:03:56 ; Search time 9.88 Seconds  
(without alignments)  
25.750 Million cell updates/sec

Title: US-08-819-669D-26

Perfect score: 52

Sequence: 1 EADPTGHSY 9

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	309	1	MAG1_HUMAN
2	44	84.6	369	1	MAG8_HUMAN
3	43	82.7	234	1	MAG8_HUMAN
4	43	82.7	315	1	MAG9_HUMAN
5	42	80.8	319	1	MAGB_HUMAN
6	37	71.2	555	1	WETA_EMENI
7	36	69.2	314	1	MAG3_HUMAN
8	36	69.2	314	1	MAG6_HUMAN
9	36	69.2	3396	1	POLG_DENIS
10	35	67.3	346	1	MG84_HUMAN
11	35	67.3	873	1	PC1_HUMAN
12	35	67.3	1033	1	TIR1_ECOLI
13	35	67.3	98	1	Y037_MYCTU
14	34	65.4	497	1	PN3_ADE12
15	34	65.4	878	1	YB9X_YEAST
16	33	63.5	1523	1	DPOL_THES9
17	33	63.5	775	1	DPOL_THES9
18	33	63.5	747	1	GUND_CELFI
19	33	63.5	2273	1	HFAL_YEAST
20	33	63.5	597	1	IXRL_YEAST
21	33	63.5	372	1	LIGC_TRAVE
22	33	63.5	1040	1	MAN1_RAT
23	33	63.5	488	1	SUOX_HUMAN
24	33	63.5	488	1	SUOX_RAT
25	33	63.5	503	1	VP57_BDV
26	32	61.5	399	1	CAQS_MOUSE
27	32	61.5	395	1	CAQS_RABIT
28	32	61.5	416	1	CC3_CANAL
29	32	61.5	669	1	COGU_HUMAN
30	32	61.5	2411	1	DAB_DROME
31	32	61.5	1002	1	DOR_DROME
32	32	61.5	700	1	EPG_AQUPY
33	32	61.5	283	1	FOLD_BAGSU
34	32	61.5	1360	1	GLI1_XENLA
35	32	61.5	1597	1	GTF1_STRDO
36	32	61.5	1592	1	GTF2_STRDO
37	32	61.5	379	1	ILEU_HORSE
38	32	61.5	280	1	LE76_BRANA
39	32	61.5	317	1	MAG4_HUMAN
40	32	61.5	349	1	NRLB_KLEPN
41	32	61.5	269	1	RAD_HUMAN
42	32	61.5	268	1	RAD_RAT
43	32	61.5	133	1	REV_CAECV

#### ALIGNMENTS

RESULT 1  
MAG1\_HUMAN

ID MAG1\_HUMAN STANDARD; PRT; 309 AA.  
AC P43355; 000346;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MELANOMA-ASSOCIATED ANTIGEN 1 (MAGE-1 ANTIGEN) (ANTIGEN M22-E).  
GN MAGE1 OR MAGE1 OR MAGE1A.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RC [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92086861.  
RA VAN DER BRUGEN P., TRAVERSARI C., CHOMEZ P., LURQUIN C., DE PLAEN E.,  
VAN DEN EYNDE B., KNUTH A., BOON T.;  
RT "A gene encoding an antigen recognized by cytolytic T lymphocytes on  
a human melanoma.";  
RL SCIENCE 254:1643-1647(1991).  
RC [2]  
RP SEQUENCE FROM N.A.  
RX TISSUE-SKIN;  
RA MEDLINE; 94311935.  
RT DING M., BECK R.J., KELLER C.J., FENTON R.G.;  
RL "Cloning and analysis of MAGE-1-related genes.";  
RN BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).  
RP SEQUENCE FROM N.A.  
RA GLOECKNER G., RUMP A., NORDSTIEK G., HINZMANN B., KIOSCHIS P.,  
HEISS N., POUSTKA A., BAUER D., DRESCHER B., KNOB A., ROSENTHAL A.;  
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
RC [4]  
RP MUTAGENESIS.  
RX TISSUE-BLOOD;  
RA MEDLINE; 94157413.  
RA GAUGLER B., VAN DEN EYNDE B., VAN DER BRUGEN P., ROMERO P.,  
GAFFORI J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;  
RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by  
autologous cytolytic T lymphocytes.";  
RL J. EXP. MED. 179:921-930(1994).  
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL  
DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR  
PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS  
CYTOLYTIC T LYMPHOCYTES.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,  
SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG,  
CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT  
FOR TESTES. NEVER EXPRESSED IN KIDNEY TUMORS, LEUKEMIAS AND  
LYMPHOMAS.  
CC -!- POLYMORPHISM: THE VARIANT AT POSITION 32 LIKELY REPRESENTS A  
POLYMORPHISM OF THE MAGE-1 GENE.  
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; M77481; G416115;  
CC EMBL; U82672; G2078527;  
CC MIN; 300016;  
CC ANTIGEN; MULTIGENE FAMILY; POLYMORPHISM; TUMOR ANTIGEN.

```
FT VARIANT 32 32 T -> A.
FT DOMAIN 33 36 POLY-SER.
FT MUTAGEN 163 163 D->A: ABOLISHES HLA-A1 BINDING.
FT MUTAGEN 169 169 Y->A: ABOLISHES HLA-A1 BINDING.
FT CONFLICT 72 72 R -> Q (IN REF. 3).
SQ SEQUENCE 309 AA: 34342 MW: E6CB1300 CRC32:

Query Match
Best Local Similarity 100.0%; Score 52; DB 1; Length 309;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
Db 161 EADPTGHSY 169

RESULT 2
MAGA_HUMAN
ID MAGA_HUMAN STANDARD; PRT; 369 AA.
AC P43363;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 10 (MAGE-10 ANTIGEN).
GN MAGEA10 OR MAGE10.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95012457.
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAENEVEE W., BOON T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL IMMUNOGENETICS 40:360-369(1994).
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG,
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U10685; G53351;
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT DOMAIN 54 62 POLY-SER.
SQ SEQUENCE 369 AA: 40766 MW: D11E1870 CRC32:

Query Match
Best Local Similarity 84.6%; Score 44; DB 1; Length 369;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
Db 193 EVDPTGHSF 201

RESULT 3
MAG8_HUMAN
ID MAG8_HUMAN STANDARD; PRT; 234 AA.
AC P43361;

Query Match
Best Local Similarity 77.8%; Score 43; DB 1; Length 234;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
Db 171 EVDPTGHSY 179

RESULT 4
MAG9_HUMAN
ID MAG9_HUMAN STANDARD; PRT; 315 AA.
AC P43362; Q92910;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 9 (MAGE-9 ANTIGEN).
GN MAGEA9 OR MAGE9.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95012457.
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAENEVEE W., BOON T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL IMMUNOGENETICS 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA TIMMS K.M., BONDESON M.L., ANSARI-LARI M.A., LAGERSTEDT K.,
```

```
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 8 (MAGE-8 ANTIGEN).
GN MAGEA8 OR MAGE8.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95012457.
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAENEVEE W., BOON T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL IMMUNOGENETICS 40:360-369(1994).
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG,
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U10693; G53352;
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT DOMAIN 40 43 POLY-SER.
SQ SEQUENCE 234 AA: 25197 MW: D4931BC3 CRC32:

Query Match
Best Local Similarity 82.7%; Score 43; DB 1; Length 234;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
Db 171 EVDPTGHSY 179

RESULT 4
MAG9_HUMAN
ID MAG9_HUMAN STANDARD; PRT; 315 AA.
AC P43362; Q92910;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 9 (MAGE-9 ANTIGEN).
GN MAGEA9 OR MAGE9.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95012457.
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAENEVEE W., BOON T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL IMMUNOGENETICS 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA TIMMS K.M., BONDESON M.L., ANSARI-LARI M.A., LAGERSTEDT K.,
```

RA NELSON D.L., PETERSSON U., GIBBS R.A.;  
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL  
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR  
CC PROGRESSION.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,  
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG  
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT  
CC FOR TESTES AND PLACENTA.  
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; U10694; G533528; -;  
DR EMBL; U66083; G1519285; -;  
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.  
FT DOMAIN 34 37 POLY-GLU.  
FT DOMAIN 87 90 POLY-GLU.  
SQ SEQUENCE 315 AA: 35088 MW; 7DC3228E CRC32;  
  
Query Match 82.7%; Score 43; DB 1; Length 315;  
Best Local Similarity 77.8%; Pred. No. 0.47;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 EADPTGHSY 9  
Db 167 EVDPAHGSY 175  
| | | | |  
| | | | |  
  
RESULT 5  
MAGB\_HUMAN  
ID MAGB\_HUMAN STANDARD; PRT; 319 AA.  
AC P43364;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MELANOMA-ASSOCIATED ANTIGEN 11 (MAGE-11 ANTIGEN).  
GN MAGE11 OR MAGE11.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95012457.  
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,  
RA DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,  
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAVEENE W., BOON T.;  
RT "Structure, chromosomal localization, and expression of 12 genes of  
RT the MAGE family".  
RL IMMUNOGENETICS 40:360-369(1994).  
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL  
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR  
CC PROGRESSION.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,  
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG  
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT  
CC FOR TESTES AND PLACENTA.  
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC  
DR EMBL; U10686; G533513; -;  
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.  
SQ SEQUENCE 319 AA; 35536 MW; E3DBDEF CRC32;  
  
Query Match 80.8%; Score 42; DB 1; Length 319;  
Best Local Similarity 77.8%; Pred. No. 0.73;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 EADPTGHSY 9  
Db 171 EVDPTSHSY 179  
| | | | |  
| | | | |  
  
RESULT 6  
WETA\_EMENI  
ID WETA\_EMENI STANDARD; PRT; 555 AA.  
AC P22022;  
DT 01-AUG-1991 (REL. 19, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE REGULATORY PROTEIN WETA.  
GN WETA.  
OS EMERICELLA NIDULANS (ASPERGILLUS NIDULANS).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PLECTOMYCETES;  
OC EUROTIALES; TRICHOCOMACEAE; EMERICELLA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91094871.  
RA MARSHALL M.A., TIMBERLAKE W.E.;  
RT "Aspergillus nidulans weta activates spore-specific gene expression".  
RL MOL. CELL. BIOL. 11:55-62(1991).  
CC -!- FUNCTION: RESPONSIBLE FOR ACTIVATING A SET OF GENES WHOSE PRODUCTS  
CC MAKE UP THE FINAL TWO CONIDIAL WALL LAYERS OR DIRECT THEIR  
CC ASSEMBLY AND THOUGH THIS ACTIVITY IS RESPONSIBLE FOR ACQUISITION  
CC OF SPORE DORMANCY.  
CC -!- FUNCTION: BRLA, ABAA & WETA ARE PIVOTAL REGULATORS OF CONIDIOPHORE  
CC DEVELOPMENT AND CONIDIUM MATURATION. THEY ACT INDIVIDUALLY AND  
CC TOGETHER TO REGULATE THEIR OWN EXPRESSION AND THAT OF NUMEROUS  
CC OTHER SPORULATION-SPECIFIC GENES.  
CC -!- DEVELOPMENTAL STAGE: THE WETA GENE IS ACTIVATED ONLY DURING  
CC CONIDIOPHORE DEVELOPMENT, AND ITS MRNA ACCUMULATES PREFERENTIALLY  
CC IN MATURE CONIDIA.  
CC -!- DOMAIN: HAS AN ACIDIC N-TERMINUS (AA 1-52) FOLLOWED BY A SER-,  
CC THR-, PRO-RICH DOMAIN (AA 125-233) AND A BASIC C-TERMINUS (AA  
CC 461-555).  
CC -!- SIMILARITY: TO P. CHRYSOGENUM WETA.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; M60528; G168109; -;  
DR PIR; A39665; RGASWA.  
KW DEVELOPMENTAL PROTEIN; CONIDIATION; TRANSCRIPTION REGULATION;  
KW ACTIVATOR.  
SQ SEQUENCE 555 AA; 60275 MW; 5B4F6B20 CRC32;  
  
Query Match 71.2%; Score 37; DB 1; Length 555;  
Best Local Similarity 87.5%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 EADPTGHS 8  
Db 109 EADATGHS 116  
| | | | |  
| | | | |

```

RESULT 7
MAG3_HUMAN
ID MAG3_HUMAN STANDARD; PRT: 314 AA.
AC P43357;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 3 (MAGE-3 ANTIGEN) (ANTIGEN M22-D).
GN MAGE3 OR MAGE3
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95012457.
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA DE SNET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAVEENE W., BOON T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family";
RL IMMUNOGENETICS 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94311935.
RA DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RL BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95369706.
RA IMAI Y., SHICHIGO S., YAMADA A., KATAYAMA T., YANO H., ITOH K.;
RT "Sequence analysis of the MAGE gene family encoding human tumor-
RT rejection antigens.";
RL GENE 160:287-290(1995).
CC -!- FUNCTION: NOT KNOWN. THOUGH MAY PLAY A ROLE IN TUMOR
CC OR ASPECTS OF TUMOR PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY TO
CC MAGE-3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U10691; G533523;
CC EMBL: U10339; G499122;
CC EMBL: D32076; G1125016;
CC ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
KW DOMAIN 40 43 POLY-SER.
SQ SEQUENCE 314 AA; 34891 MW; B7125E97 CRC32;

Query Match 69.2%; Score 36; DB 1; Length 314;
Best Local Similarity 66.7%; Pred. NO. 9.8;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
DB 168 EVDPIGHVY 176

RESULT 8
MAG6_HUMAN
ID MAG6_HUMAN STANDARD; PRT: 314 AA.
AC P43360;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 6 (MAGE-6 ANTIGEN) (MAGE3B).
GN MAGE6 OR MAGE6
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94157413.
RA GAUGLER B., VAN DEN EYNDE B., VAN DER BRUGGEN P., ROMERO P.,
RA GAFORIO J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;
RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
RT autologous cytolytic T lymphocytes.";
RL J. EXP. MED. 179:921-930(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94311935.
RA DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RL BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
CC -!- FUNCTION: NOT KNOWN. THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC CYTOLYTIC T LYMPHOCYTES.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA. NEVER EXPRESSED IN KIDNEY TUMORS,
CC LEUKEMIAS AND LYMPHOMAS.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U03735; G468826;
CC ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
KW DOMAIN 40 43 POLY-SER.
FT MUTAGEN 170 170 D->A: ABOLISHES HLA-A1 BINDING.
FT MUTAGEN 176 176 Y->A: ABOLISHES HLA-A1 BINDING.
SQ SEQUENCE 314 AA; 34747 MW; AC557A64 CRC32;

Query Match 69.2%; Score 36; DB 1; Length 314;
Best Local Similarity 66.7%; Pred. NO. 9.8;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
DB 168 EVDPIGHVY 176

RESULT 9
POLG_DENIS
ID POLG_DENIS STANDARD; PRT: 3396 AA.
AC P33478;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DE GENOME POLYPEPTIDE [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX
DE PROTEIN (ENVELOPE PROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL
DE PROTEIN NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3); RNA-DIRECTED
DE RNA POLYMERASE (EC 2.7.7.48) (NS5)].
OS DENGUE VIRUS TYPE 1 (STRAIN SINGAPORE S275/90).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;

```

OC FLAVIVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92263809.  
RA FU J., TAN B.H., YAP E.H., CHAN Y.C., TAN Y.H.;  
RT "Full-length cDNA sequence of dengue type 1 virus (Singapore strain  
RT S275/90)."  
RL VIROLOGY 188:953-958(1992).  
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF THREE PROTEINS:  
CC PROTEIN PM, PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A  
CC COMPLEX OF PROTEIN C AND MRNA.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M87512; -; NOT\_ANNOTATED\_CDS.  
CC  
CC PIR; A42551; A42551.  
CC PFAM; PF00869; Flavi\_glycoprot; 1.  
CC PFAM; PF00948; Flavi\_Ns1; 1.  
CC PFAM; PF00949; Flavi\_helicase; 1.  
CC PFAM; PF00972; Flavi\_Ns2; 1.  
CC PFAM; PF01002; Flavi\_Ns2B; 1.  
CC PFAM; PF01003; Flavi\_capsid; 1.  
CC PFAM; PF01004; Flavi\_M; 1.  
CC PFAM; PF01005; Flavi\_Ns2A; 1.  
CC HSP; P14336; ISVB.  
CC  
CC POLYPROTEIN; GLYCOPROTEIN; TRANSFERASE; RNA-DIRECTED RNA POLYMERASE;  
CC CORE PROTEIN; COAT PROTEIN; ENVELOPE PROTEIN; HELICASE; ATP-BINDING;  
CC TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.  
CC  
CC FT CHAIN 1 114  
CC FT PROPEP 115 205  
CC FT CHAIN 206 280  
CC FT CHAIN 281 774  
CC FT CHAIN 775 1127  
CC FT CHAIN 1128 1344  
CC FT CHAIN 1345 1474  
CC FT CHAIN 1475 2093  
CC FT CHAIN 2094 2243  
CC FT CHAIN 2244 2492  
CC FT CHAIN 2493 3396  
CC  
CC NP\_BIND 1668 1675  
CC SITE 1759 1762  
CC TRANSMEM 267 279  
CC TRANSMEM 753 769  
CC FT DISULFID 283 310  
CC FT DISULFID 340 396  
CC FT DISULFID 354 385  
CC FT DISULFID 372 401  
CC FT DISULFID 465 565  
CC FT DISULFID 582 613  
CC FT CARBOHYD 183 183  
CC FT CARBOHYD 347 347  
CC FT CARBOHYD 433 433  
CC SEQUENCE 3396 AA; 379558 MW; F538880D CRC32;

Query Match 69.2%; Score 36; DB 1; Length 3396;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EADPTGHS 8  
Db 3383 ESDPKGHS 3390  
RESULT 10

MG84\_HUMAN  
ID MG84\_HUMAN STANDARD; PRT; 346 AA.  
AC O15481;  
DT 15-DEC-1998 (REL. 37, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE MELANOMA-ASSOCIATED ANTIGEN B4 (MAGE-B4 ANTIGEN).  
GN MAGE-B4.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98110375.  
RA LURQUIN C., DE SMT C., BRASSEUR F., MUSCATELLI F., MARTELANGE V.,  
RA DE PLAEN E., BRASSEUR R., MONACO A.P., BOON T.;  
RT "Two members of the human MAGEB gene family located in Xp21.3 are  
RT expressed in tumors of various histological origins."  
RL GENOMICS 46:397-408(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MUZY D., ARENSON A.D., ADAMS C., BRUNDAGE E., BUNAC C., CARVELLI K.,  
RA CHACKO J., CHEN J., DI W., DING Y., DUCAN S., DURBIN J., FORCUM J.,  
RA GANESH R., GARCIA C., GOODMAN M., GORRELL J.H., HAYWOOD M.,  
RA HERNANDEZ J., JACKSON L., JIN S., KAMPAL R., KARPATY S., KOVAR C.,  
RA LEAL B., LI Y., LICHTARGE O., LIU W., LOGAN O., LU J., LY T.,  
RA MARTINEZ C., OSWAL G., PEREZ L., RASHID N.D., ROWLAND K., SAVAGE L.,  
RA SCHERER S.E., SHEN H., SIMON M., STOVALL K., TIMMS K.M., TODD J.,  
RA VO O., WILLIAMSON A., WORLEY K.C., YU W., CHINAULT C., NELSON D.,  
RA GIBBS R.A.;  
RL SUBMITTED (OCT-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN TESTIS.  
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U93163; G2459682; -.  
CC EMBL; AC005185; G3687199; -.  
CC MIM; 300153; -.  
CC KW ANTIGEN; MULTIGENE FAMILY.  
SQ SEQUENCE 346 AA; 38923 MW; C6A5A407 CRC32;  
Query Match 67.3%; Score 35; DB 1; Length 346;  
Best Local Similarity 66.7%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EADPTGHS 9  
Db 168 EVNPTGHSY 176  
RESULT 11  
PC1\_HUMAN  
ID PC1\_HUMAN STANDARD; PRT; 873 AA.  
AC P22413;  
DT 01-AUG-1991 (REL. 19, CREATED)  
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 (ALKALINE PHOSPHODIESTERASE I  
DE (EC 3.1.4.1) / NUCLEOTIDE PYROPHOSPHATASE (EC 3.6.1.9) (NPPASE)).  
GN PC1.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE; 91009202.  
RA BUCKLEY M.F., LOVELAND K.A., MCKINSTRY W.J., GARSON O.M., GODING J.W.;  
RT "Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human  
RT molecule, amino acid sequence, and chromosomal location.";  
RL J. BIOL. CHEM. 265:17508-17511(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92246539.  
RA FUNAKOSHI I., KATO H., HORIE K., YANO T., HORI Y., KOBAYASHI H.,  
RA INOUE T., SUZUKI H., FUKUI S., TSUKAHARA M., KAJI T.,  
RA YAMASHINA I.;  
RT "Molecular cloning of cDNAs for human fibroblast nucleotide  
RT pyrophosphatase.";  
RL ARCH. BIOCHEM. 295:180-187(1992).  
CC -!- FUNCTION: MAY HAVE A ROLE IN THE REGULATION OF N-GLYCOSYLATION.  
CC -!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES  
CC SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED  
CC OLIGO-NUCLEOTIDES.  
CC -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PLASMA CELLS AND ALSO IN A NUMBER  
CC OF NON-LYMPHOID TISSUES, INCLUDING THE DISTAL CONVOLUTED TUBULE  
CC OF THE KIDNEY, CHONDROCYTES, AND EPIDIDYMIS.  
CC -!- SIMILARITY: CONTAINS 2 SOMATOMEDIN-B TYPE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; M57736; G189650; -  
DR EMBL; D12485; G219945; -  
DR EMBL; D12485; G219944; ALT\_INIT.  
DR PIR; A39216; A39216.  
DR MIM; 173335; -  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
DR PFAM; PF01033; Somatomedin\_B; 2.  
KW GLYCOPROTEIN; TRANSMEMBRANE; DUPLICATION; SIGNAL-ANCHOR; HYDROLASE.  
FT DOMAIN 1 24  
FT TRANSMEM 25 45  
FT DOMAIN 46 873  
FT DOMAIN 52 92  
FT DOMAIN 93 136  
FT CARBOHYD 127 127  
FT CARBOHYD 233 233  
FT CARBOHYD 289 289  
FT CARBOHYD 425 425  
FT CARBOHYD 533 533  
FT CARBOHYD 591 591  
FT CARBOHYD 648 648  
FT CARBOHYD 679 679  
FT CARBOHYD 696 696  
SQ SEQUENCE 873 AA; 99929 MW; 580583CD CRC32;  
  
Query Match 67.3%; Score 35; DB 1; Length 873;  
Best Local Similarity 66.7%; Pred. NO. 44;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 EADPTGHSY 9  
DB 322 EPDSSGHSY 330  
  
RESULT 12  
T1RL\_ECOLI  
ID T1RL\_ECOLI STANDARD; PRT; 1033 AA.  
AC P10486;  
DT 01-JUL-1989 (REL. 11, CREATED)

DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE TYPE I RESTRICTION ENZYME ECOL124II R PROTEIN (EC 3.1.21.3).  
GN HSR OR HSR.  
OS ESCHERICHIA COLI.  
OG PLASMID INCFIV R124/3.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
OC ESCHERICHIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89178628.  
RA PRICE C., LINGNER J., BICKLE J., FIRMAN T.A., GLOVER S.W.;  
RT "Basis for changes in DNA recognition by the Ecor124 and Ecor124/3  
RT type I DNA restriction and modification enzymes.";  
RL J. MOL. BIOL. 205:115-125(1989).  
CC -!- FUNCTION: THE Ecor124/3 I ENZYME RECOGNIZES 5'GAA(N7)RTCG.  
CC -!- FUNCTION: SUBUNIT R IS REQUIRED FOR BOTH NUCLEASE AND ATPASE  
CC ACTIVITIES, BUT NOT FOR MODIFICATION.  
CC -!- SUBUNIT: THE TYPE I RESTRICTION & MODIFICATION SYSTEM IS COMPOSED  
CC OF THREE POLYPEPTIDES R, M AND S.  
CC -!- TYPE I RESTRICTION AND MODIFICATION ENZYMES ARE COMPLEX. MULTI-  
CC FUNCTIONAL SYSTEMS WHICH REQUIRE ATP, S-ADENOSYL METHIONINE AND  
CC MG(2+) AS CO-FACTORS AND, IN ADDITION TO THEIR ENDONUCLEOLYTIC  
CC AND METHYLASE ACTIVITIES, ARE POTENT DNA-DEPENDENT ATPASES.  
CC -!- SIMILARITY: WITH ATPASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X13145; G41750; -  
DR PIR; S02168; S02168.  
DR REBASE; RB00748; Ecor124II.  
KW PLASMID; RESTRICTION SYSTEM; HYDROLASE; DNA-BINDING; ATP-BINDING.  
SQ SEQUENCE 1033 AA; 119656 MW; 9E988CC1 CRC32;  
  
Query Match 67.3%; Score 35; DB 1; Length 1033;  
Best Local Similarity 66.7%; Pred. NO. 52;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 EADPTGHSY 9  
DB 22 KAEPTGDSY 30  
  
RESULT 13  
Y037\_MYCTU  
ID Y037\_MYCTU STANDARD; PRT; 98 AA.  
AC Q10635;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE VERY HYPOTHETICAL 11.3 KD PROTEIN CY130.07.  
GN MYC130.07.  
OS MYCOBACTERIUM TUBERCULOSIS.  
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIACEAE;  
OC ACTINOMYCETALES; CORYNEBACTERIACEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; 273902; E245016; -  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 98 AA; 11334 MW; 8E3C8C75 CRC32;

Query Match 67.3%; Score 35; DB 1; Length 98;  
Best Local Similarity 66.7%; Pred. No. 4.5;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9  
| | | | |  
Db 24 EAGPDGHEY 32

RESULT 14  
PEN3\_ADEL2  
ID PEN3\_ADEL2 STANDARD; PRT; 497 AA.  
AC P36716;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PENTON PROTEIN (VIRION COMPONENT III) (PENTON BASE PROTEIN).  
GN PIII.  
OS HUMAN ADENOVIRUS TYPE 12.  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; ADENOVIRIDAE; MASTADENOVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94076430.

RA SPRENGEL J., SCHMITZ B., HEUSS-NEITZEL D., ZOCK C., DOERFLER W.;  
RT "Nucleotide sequence of human adenovirus type 12 DNA: comparative  
functional analysis."  
RL J. VIROL. 68:379-389(1994).  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; X73487; G313372; -  
DR PIR; S33938; S33938.  
KW LATE PROTEIN.  
SQ SEQUENCE 497 AA; 56393 MW; A5BEC571 CRC32;

Query Match 65.4%; Score 34; DB 1; Length 497;  
Best Local Similarity 66.7%; Pred. No. 38;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9  
| | | | |  
Db 310 ETDPKGRSY 318

RESULT 15  
YB9X\_YEAST  
ID YB9X\_YEAST STANDARD; PRT; 878 AA.  
AC P38149;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 98.1 KD TRP-ASP REPEATS CONTAINING PROTEIN IN PAF1-MRPL27  
DE INTERGENIC REGION.  
GN YBR281C OR YBR2018.  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-S286C;  
RX MEDLINE; 94378722.  
RA HOLMSTROEM K., BRANDT T., KALLESOE T.;  
RT "The sequence of a 32,420 bp segment located on the right arm of  
RT chromosome II from Saccharomyces cerevisiae."  
RL YEAST 10:47-62(1994).  
CC -!- SIMILARITY: CONTAINS ? WD REPEATS (TRP-ASP DOMAINS).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; X76053; G429126; -  
DR EMBL; Z36150; G536726; -  
DR PIR; S44543; S44543.  
DR PIR; S39137; S39137.  
DR PROSITE; PS00678; WD\_REPEATS; 2.  
KW HYPOTHETICAL PROTEIN; REPEAT; WD REPEAT.  
SQ SEQUENCE 878 AA; 98070 MW; 093B13F4 CRC32;

Query Match 65.4%; Score 34; DB 1; Length 878;  
Best Local Similarity 71.4%; Pred. No. 68;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 DPTGHSY 9  
| | | | |  
Db 796 DPTNHAY 802

Search completed: December 8, 1999, 23:56:43  
Job time: 17567 sec





GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 1999, 00:13:00 ; Search time 15.25 Seconds  
(without alignments)  
36.321 Million cell updates/sec

Title: US-08-819-669D-26

Perfect score: 52

Sequence: 1 EADPTGHSY 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	71.2	1032	11 Q61989	Q61989 mus musculus
2	36	69.2	2457	12 Q41965	Q41965 murine herp
3	35	67.3	1187	2 Q59278	Q59278 cellulomona
4	35	67.3	131	2 Q85701	Q85701 streptomyc
5	35	67.3	330	11 Q60763	Q60763 mus musculu
6	35	67.3	330	11 Q60761	Q60761 mus musculu
7	35	67.3	320	11 Q89006	Q89006 mus musculu
8	35	67.3	325	11 Q89010	Q89010 mus musculu
9	34	65.4	1671	1 P77933	P77933 pyrococcus
10	34	65.4	197	2 Q53701	Q53701 mycobacteri
11	34	65.4	224	2 Q69955	Q69955 streptomyc
12	34	65.4	215	2 Q87835	Q87835 streptomyc
13	34	65.4	268	4 Q14969	Q14969 homo sapien
14	34	65.4	1184	4 Q75339	Q75339 homo sapien
15	34	65.4	347	4 Q75862	Q75862 homo sapien
16	34	65.4	347	4 Q00601	Q00601 homo sapien
17	34	65.4	156	5 Q23961	Q23961 drosophila
18	34	65.4	599	6 Q19112	Q19112 sus scrofa
19	34	65.4	272	11 Q35830	Q35830 rhizomys pr
20	34	65.4	267	11 Q70217	Q70217 mus musculu
21	34	65.4	342	11 Q88737	Q88737 mus musculu
22	34	65.4	320	11 Q89009	Q89009 mus musculu
23	34	65.4	906	12 Q9YU1	Q9YU1 hemorhagic
24	34	65.4	370	13 Q91904	Q91904 xenopus lae
25	33	63.5	295	1 Q26704	Q26704 methanobact
26	33	63.5	301	1 Q59491	Q59491 pyrococcus
27	33	63.5	1086	2 Q69230	Q69230 bacillus sp
28	33	63.5	307	2 Q24828	Q24828 acinetobact
29	33	63.5	288	2 Q51323	Q51323 pseudomonas

## ALIGNMENTS

### RESULT 1

Q61989 ID Q61989 PRELIMINARY; PRT; 1032 AA.  
AC Q61989; Q61740;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
DE INTEGRIN ALPHA-4 SUBUNIT PRECURSOR (INTEGRIN ALPHA-IV) (VLA-4) (CD49D)  
DE (LYMPHOCYTE-PEYER'S PATCH ADHESION MOLECULES ALPHA SUBUNIT) (LPAM)  
DE ALPHA SUBUNIT)  
GN ITGA4 OR VLA-4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RC STRAIN=OBA, AND NIH/SWISS; TISSUE=LEUKEMIA, AND FIBROBLAST;  
RX MEDLINE: 95290094.  
RA DE MEIRSMAN C., SCHOLLEN E., JASPERS M., ONGENA K., MATTHIJS G.,  
RA MARYNEN P., CASSIMAN J.J.;  
RT "Cloning and characterization of the promoter region of the murine  
alpha-4 integrin subunit.";  
RL DNA CELL BIOL. 13:743-754(1994).  
RN [2]  
SEQUENCE FROM N.A.  
RC STRAIN=B6/CBA, AND NIH/SWISS; TISSUE=SPLEEN, AND FIBROBLAST;  
RX MEDLINE: 96326295.  
RA DE MEIRSMAN C., JASPERS M., SCHOLLEN E., CASSIMAN J.J.;  
RL "The genomic structure of the murine alpha 4 integrin gene.";  
RL DNA CELL BIOL. 15:595-603(1996).  
RN [3]  
SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=LIVER;  
RA HASEGAWA M., FOOTE S.;  
RL SUBMITTED (NOV-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [4]  
SEQUENCE OF 134-307 FROM N.A.  
TISSUE=MELANOMA;  
RA ROUT U.K., ARMANI D.R.;  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: THE ALPHA-4 BETA-1 INTEGRIN MEDIATES ADHESION TO ONE OR  
MORE DOMAINS WITHIN THE HEP1-11CS REGION OF FIBRONECTIN  
INCLUDING THE ALTERNATIVELY SPLICED CS-1 REGION. VLA-4 ALSO  
RECOGNIZES THE CELLULAR LIGAND VCAM-1 ON ACTIVATED ENDOTHELIAL  
CELLS. IT TRIGGERS HOMOTYPIC AGGREGATION FOR MOST VLA-4-POSITIVE  
LEUKOCYTE CELL LINES. VLA-4 MAY ALSO PARTICIPATE IN CYTOLYTIC  
T-CELL INTERACTIONS WITH TARGET CELLS. IMPORTANT FOR CELL-CELL  
ADHESION FUNCTIONS (BY SIMILARITY).  
CC -!- FUNCTION: THE ALPHA-4 BETA-7 HETERODIMER INTERACTS WITH THE  
LYMPHOCYTE-PEYER'S PATCH ADHESION MOLECULE MADCAM-1 AND IS  
IMPORTANT FOR CELL-CELL ADHESION FUNCTIONS. VLA-4 MAY PLAY A  
SUPPORTIVE ROLE IN THE HOMING OF LYMPHOCYTES TO THE PEYER'S

30 33 63.5 2273 3 Q42823 042823 saccharomyc  
31 33 63.5 827 3 Q43048 043048 schizosacch  
32 33 63.5 341 5 Q25028 025028 haemochus  
33 33 63.5 341 10 Q43449 043449 glycine max  
34 33 63.5 245 10 Q23713 023713 arabidopsis  
35 33 63.5 246 10 Q81146 081146 arabidopsis  
36 33 63.5 246 10 Q81147 081147 arabidopsis  
37 33 63.5 106 10 Q92RP0 092RP0 caragana ar  
38 33 63.5 503 12 Q89857 089857 borna disea  
39 33 63.5 503 12 Q65459 065459 borna disea  
40 33 63.5 503 12 Q10394 010394 borna disea  
41 33 63.5 503 12 Q10397 010397 borna disea  
42 33 63.5 503 12 Q10399 010399 borna disea  
43 33 63.5 503 12 Q10400 010400 borna disea  
44 33 63.5 503 12 Q10403 010403 borna disea  
45 33 63.5 353 12 Q88626 088626 borna disea

CC PATCHES.  
 CC -!- SUBUNIT: DIMER OF AN ALPHA AND BETA CHAIN. THE BETA CHAIN THAT  
 CC ASSOCIATES WITH ALPHA-4 IS EITHER BETA-1 OR BETA-7. THE ALPHA-4  
 CC BETA-1 HETERODIMER IS REFERRED TO AS VLA-4 OR LPAM-2. THE ALPHA-4  
 CC BETA-7 HETERODIMER IS REFERRED TO AS LPAM-1 OR THE PEYER'S PATCH  
 CC HOMING RECEPTOR. THE ALPHA CHAIN CAN SOMETIMES BE CLEAVED INTO TWO  
 CC NON-COVALENTLY ASSOCIATED FRAGMENTS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: THE VLA-4 INTEGRIN IS MAINLY EXPRESSED BY  
 CC LYMPHOID AND MYELOID CELLS. THE ALPHA-4 BETA-7 INTEGRIN IS  
 CC EXPRESSED BY PEYER'S PATCH HOMING CELLS.  
 CC -!- MISCELLANEOUS: THIS PROTEIN HAS 3 POTENTIAL CALCIUM-BINDING SITES.  
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.  
 DR EMBL: L20788; AAA97501.1; -  
 DR EMBL: U34800; AAB09630.1; -  
 DR EMBL: U34827; AAB09630.1; JOINED.  
 DR EMBL: U34828; AAB09630.1; JOINED.  
 DR EMBL: U34829; AAB09630.1; JOINED.  
 DR EMBL: U34830; AAB09630.1; JOINED.  
 DR EMBL: U34831; AAB09630.1; JOINED.  
 DR EMBL: U34832; AAB09630.1; JOINED.  
 DR EMBL: U34833; AAB09630.1; JOINED.  
 DR EMBL: U34834; AAB09630.1; JOINED.  
 DR EMBL: U34835; AAB09630.1; JOINED.  
 DR EMBL: U34763; AAB09630.1; JOINED.  
 DR EMBL: U34764; AAB09630.1; JOINED.  
 DR EMBL: U34765; AAB09630.1; JOINED.  
 DR EMBL: U34766; AAB09630.1; JOINED.  
 DR EMBL: U34767; AAB09630.1; JOINED.  
 DR EMBL: U34768; AAB09630.1; JOINED.  
 DR EMBL: U34769; AAB09630.1; JOINED.  
 DR EMBL: U34770; AAB09630.1; JOINED.  
 DR EMBL: U34797; AAB09630.1; JOINED.  
 DR EMBL: U34798; AAB09630.1; JOINED.  
 DR EMBL: U34799; AAB09630.1; JOINED.  
 DR EMBL: AF109136; AAC95388.1; -  
 DR MGD; MGI:96603; ITGA4.  
 DR PFAM; PF00357; Integrin.A.1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 KW Integrin; Cell adhesion; Glycoprotein; Transmembrane; Signal;  
 KW Extracellular matrix; Cytoskeleton; Calcium-binding; Polymorphism.  
 FT SIGNAL 1 33 POTENTIAL.  
 FT CHAIN 34 1032 INTEGRIN ALPHA-4 SUBUNIT.  
 SQ SEQUENCE 1032 AA; 115085 MW; 27B379B1 CRC32;

Query Match 71.2%; Score 37; DB 11; Length 1032;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PTGHSY 9  
 DB 29 PTGHSY 34

RESULT 2  
 O41965 PRELIMINARY; PRT; 2457 AA.  
 ID O41965;  
 AC O41965;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE TEGUMENT PROTEIN.  
 GN GAMMAHV. ORF64.  
 OS murine herpesvirus 68.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WUMS;  
 RX MEDLINE; 97366649.

RA VIRGIN H.W. IV, LATREILLE P., WAMSLEY P., HALLSWORTH K., WECK K.E.,  
 RA DAL CANTO A.J., SPECK S.H.;  
 RT "Complete sequence and genomic analysis of murine gammaherpesvirus  
 68";  
 RL J. Virol. 71:5894-5904(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WUMS;  
 RA LATREILLE P., WAMSLEY P., WATERSTON R.H.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U97553; AAB66417.1; -  
 SQ SEQUENCE 2457 AA; 273535 MW; 8DC81BBD CRC32;

Query Match 69.2%; Score 36; DB 12; Length 2457;  
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DPTGHSY 9  
 DB 1531 DPTGHSY 1537

RESULT 3  
 Q59278 PRELIMINARY; PRT; 1187 AA.  
 ID Q59278;  
 AC Q59278;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE ENDOXYLANASE (EC 3.2.1.8) (ENDO-1,4-BETA-XYLANASE)  
 DE (1,4-BETA-D-XYLAN XYLANOXYDROLASE).  
 GN XYN.  
 OS Cellulomonas.fim.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Micrococineae; Cellulomonadaceae; Cellulomonas.  
 RN [1]  
 RP SEQUENCE OF 1-352 FROM N.A.  
 RX MEDLINE; 96245431.  
 RA CLARKE J.H.; DAVIDSON K., GILBERT H.J., FONTES C.M., HAZLEWOOD G.P.;  
 RT "A modular xylanase from mesophilic Cellulomonas fiml contains the  
 RT same cellulose-binding and thermostabilizing domains as xylanases  
 RT from thermophilic bacteria";  
 RL FEMS Microbiol. Lett. 139:27-35(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA CLARKE J.H.;

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC  
 CC LINKAGES IN XYLANS.  
 DR EMBL; Z50866; CAA90745.1; -  
 DR PFAM; PF00331; Glyco\_hydro\_10; 1.  
 KW Xylan degradation; Hydrolase; Glycosidase.  
 SQ SEQUENCE 1187 AA; 125378 MW; 92B3994A CRC32;

Query Match 67.3%; Score 35; DB 2; Length 1187;  
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 DPTGHSY 9  
 DB 1047 DPTGHSY 1053

RESULT 4  
 O85701 PRELIMINARY; PRT; 131 AA.  
 ID O85701;  
 AC O85701;  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL 14.2 KD PROTEIN.

OS Streptomyces lividans.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1326, AJ100;  
 RA ALTENBUCHNER J.;  
 RT "Amplifiable element AUD4 from Streptomyces lividans 66.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF072709; AAC35770.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 131 AA; 14187 MW; 8321BCE1 CRC32;

Query Match 67.3%; Score 35; DB 2; Length 131;  
 Best Local Similarity 62.5%; Pred. No. 12;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ADPTGHSY 9

Db 110 SDPAGHSF 117

RESULT 5  
 Q60763 PRELIMINARY; PRT; 330 AA.  
 AC Q60763;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE MELANOMA ANTIGEN, RELATED SEQUENCE 2 (SMAGE-3 PROTEIN).  
 GN MAGE-RS3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DBA/2; TISSUE=KIDNEY;  
 RX MEDLINE; 96070435.  
 RA DE BACKER O., VERHEYDEN A.M., MARTIN B., GODELAINE D., DE PLAEN E.,  
 RA BRASSEUR R., AVNER P., BOON T.;  
 RT "Structure, chromosomal location, and expression pattern of three  
 mouse genes homologous to the human MAGE genes.";  
 RL Genomics 28:74-83(1995).  
 DR EMBL; U19033; AAA86098.1;  
 DR MGD; MGI:105109; MAGE-RS3.  
 DR PFAM; PF01454; MAGE; 1.  
 SQ SEQUENCE 330 AA; 35985 MW; 83AD4246 CRC32;

Query Match 67.3%; Score 35; DB 11; Length 330;  
 Best Local Similarity 66.7%; Pred. No. 31;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9

Db 161 EIDPSTHSY 169

RESULT 6  
 Q60761 PRELIMINARY; PRT; 330 AA.  
 AC Q60761; Q60762;  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE MELANOMA ANTIGEN RELATED SEQUENCES 1 AND 2  
 (SMAGE-1 PROTEIN / SMAGE-2 PROTEIN).  
 GN (SMAGE-RS1 OR SMAGE1) AND (SMAGE-RS2 OR SMAGE2).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=DBA/2; TISSUE=KIDNEY;  
 RX MEDLINE; 96070435.  
 RA DE BACKER O., VERHEYDEN A.M., MARTIN B., GODELAINE D., DE PLAEN E.,  
 RA BRASSEUR R., AVNER P., BOON T.;  
 RT "Structure, chromosomal location, and expression pattern of three  
 mouse genes homologous to the human MAGE genes.";  
 RL Genomics 28:74-83(1995).  
 CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.  
 CC TYPES BUT NOT IN NORMAL TISSUES EXCEPT TESTIS.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN TUMOURS OF VARIOUS HISTOLOGICAL  
 DR EMBL; U19031; AAA86096.1; ALT\_INIT.  
 DR EMBL; U19032; AAA86097.1;  
 DR MGD; MGI:105117; MAGE-RS2.  
 DR PFAM; PF01454; MAGE; 1.  
 KW Antigen; Tumor antigen.  
 SQ SEQUENCE 330 AA; 35936 MW; 36D760C5 CRC32;

Query Match 67.3%; Score 35; DB 11; Length 330;  
 Best Local Similarity 66.7%; Pred. No. 31;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9

Db 161 EIDPSTHSY 169

RESULT 7  
 Q89006 PRELIMINARY; PRT; 320 AA.  
 ID Q89006;  
 AC Q89006;  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE MAGEAL PROTEIN.  
 GN MAGEAL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA DE PLAEN E., DE BACKER O., ARNAUD D., BONJEAN B., CHOMEZ P.,  
 RA MARTELANGE V., AVNER P., BALDACCIO P., BABINET C., HWANG SY.,  
 RA KNOWLES B., BOON T.;  
 RT "A new family of mouse genes homologous to the human MAGE genes.";  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ005525; CAA06579.1;  
 DR PFAM; PF01454; MAGE; 1.  
 SQ SEQUENCE 320 AA; 36096 MW; 914E8580 CRC32;

Query Match 67.3%; Score 35; DB 11; Length 320;  
 Best Local Similarity 66.7%; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9

Db 155 EVDPSVHSY 163

RESULT 8  
 Q89010 PRELIMINARY; PRT; 325 AA.  
 ID Q89010;  
 AC Q89010;  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE MAGEA6 PROTEIN.  
 GN MAGEA6.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA DE PLAEN E., DE BACKER O., ARNAUD D., BONJEAN B., CHOMEZ P.,  
 RA MARTELANGE V., AVNER P., BALDACCI P., BABINET C., HWANG SY.,  
 RA KNOWLES B., BOON T.;  
 RT "A new family of mouse genes homologous to the human MAGE genes.";  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ005530; CAA06384.1;  
 DR PFAM: PF01454; MAGE.1;  
 SQ SEQUENCE 325 AA; 36640 MW; 3B3022A3 CRC32;  
 Query Match 67.3%; Score 35; DB 11; Length 325;  
 Best Local Similarity 66.7%; Pred. No. 31;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EADPTGHSY 9  
 I | | | | |  
 Db 155 EVDPVSHSY 163  
 RESULT 9  
 P77933 PRELIMINARY; PRT; 1671 AA.  
 AC P77933;  
 DT 01-FEB-1997 (TEMBLrel.. 02, Created)  
 DT 01-JAN-1998 (TEMBLrel.. 05, Last sequence update)  
 DT 01-NOV-1998 (TEMBLrel.. 08, Last annotation update)  
 DE DNA-DEPENDENT DNA POLYMERASE (EC 2.7.7.7)  
 DE (DNA-DIRECTED DNA POLYMERASE) (DNA NUCLEOTIDYLTRANSFERASE  
 DE (DNA-DIRECTED))  
 GN POL.  
 OS Pyrococcus sp.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KOD1;  
 RX MEDLINE; 98027387.  
 RA TAKAGI M., NISHIOKA M., KAKIHARA H., KITABAYASHI M., INOUE H.,  
 RA KAKAGAMI B., OKA M., IMAKAWA T.;  
 RT "Characterization of DNA polymerase from Pyrococcus sp. strain KOD1  
 RT and its application to PCR.";  
 RL Appl. Environ. Microbiol. 63:4504-4510(1997).  
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N  
 CC DIPHOSPHATE + DNA(N).  
 DR EMBL: D29671; BAA06142.1;  
 DR PFAM: PF00136; DNA\_pol\_B; 3.  
 KW Transferase; Nucleotidyltransferase; Endonuclease.  
 FT CHAIN 407 766  
 FT ENDONUCLEASE.  
 FT CHAIN 852 1388  
 FT ENDONUCLEASE.  
 SQ SEQUENCE 1671 AA; 193432 MW; 33C27AEE CRC32;  
 Query Match 65.4%; Score 34; DB 1; Length 1671;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EADPTGHSY 9  
 I | | | | |  
 Db 1616 EFDPTKHY 1624  
 RESULT 10  
 O53701 PRELIMINARY; PRT; 197 AA.  
 AC O53701;  
 DT 01-JUN-1998 (TEMBLrel.. 06, Created)  
 DT 01-JUN-1998 (TEMBLrel.. 06, Last sequence update)  
 DT 01-NOV-1998 (TEMBLrel.. 08, Last annotation update)  
 DE HYPOTHEICAL 21.4 KD PROTEIN.  
 GN MT036.01C.

OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RA BROWN D., CHURCHER C.M.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RA COLE S.T., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE; 96181548.  
 RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,  
 RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,  
 RA COLE S.T.;  
 RT "An integrated map of the genome of the tubercle bacillus,  
 RT Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium  
 RT lepreae.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).  
 DR EMBL: AL021931; CAA17372.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 197 AA; 21376 MW; FABC6607 CRC32;  
 Query Match 65.4%; Score 34; DB 2; Length 197;  
 Best Local Similarity 71.4%; Pred. No. 28;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 DPTGHSY 9  
 I | | | | |  
 Db 46 DPTSHAY 52  
 RESULT 11  
 O69955 PRELIMINARY; PRT; 224 AA.  
 AC O69955;  
 DT 01-AUG-1998 (TEMBLrel.. 07, Created)  
 DT 01-AUG-1998 (TEMBLrel.. 07, Last sequence update)  
 DT 01-NOV-1998 (TEMBLrel.. 08, Last annotation update)  
 DE HYPOTHEICAL 24.1 KD PROTEIN (FRAGMENT).  
 GN SC3F9.15.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA SEEGER K.J., HARRIS D.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE; 97000351.  
 RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,  
 RA KINASHI H., HOPWOOD D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL023862; CAA19638.1;  
 KW Hypothetical protein.  
 FT NON\_TER 224 224  
 SQ SEQUENCE 224 AA; 24075 MW; 75A2A272 CRC32;

Query Match 65.4%; Score 34; DB 2; Length 224;  
 Best Local Similarity 85.7%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 DPTGHSY 9  
 DB 204 DPTGSSY 210

RESULT 12  
 O87835 PRELIMINARY; PRT; 215 AA.  
 AC O87835;  
 DT 01-NOV-1998 (TREMELrel. 08, Created)  
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)  
 DE HYPOTHETICAL 24.2 KD PROTEIN (FRAGMENT).  
 GN SC8A6.01.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA SEGER K.J., HARRIS D.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE; 97000351.  
 RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,  
 RA KINASHI H., HOPWOOD D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL031013; CAA19773.1; -;  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 FT NON\_TER 215 215  
 SQ SEQUENCE 215 AA; 24196 MW; F7643103 CRC32;

Query Match 65.4%; Score 34; DB 2; Length 215;  
 Best Local Similarity 85.7%; Pred. No. 31;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 DPTGHSY 9  
 DB 15 DPTGSSY 21

RESULT 13  
 Q14969 PRELIMINARY; PRT; 268 AA.  
 AC Q14969;  
 DT 01-NOV-1996 (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)  
 DE T(3:5)(Q25.1:P34) FUSION GENE NPM-MLF1.  
 GN NPM-MLF1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96152893.

RA YONEDA-KATO N., LOOK A.T., KIRSTEIN M.N., VALENTINE M.B.,  
 RA RAIMONDI S.C., COHEN K.J., CARROLL A.J., MORRIS S.W.;  
 RT "The t(3;5)(q25.1;q34) of myelodysplastic syndrome and acute myeloid  
 RT leukemia produces a novel fusion gene, NPM-MLF1.";  
 RL Oncogene 12:265-275(1996).  
 DR EMBL; L49054; AAA99997.1; -;  
 SQ SEQUENCE 268 AA; 30627 MW; F1B19980 CRC32;

Query Match 65.4%; Score 34; DB 4; Length 268;  
 Best Local Similarity 71.4%; Pred. No. 39;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DPTGHSY 9  
 DB 99 DPNHSHF 105

RESULT 14  
 O75339 PRELIMINARY; PRT; 1184 AA.  
 AC O75339;  
 DT 01-NOV-1998 (TREMELrel. 08, Created)  
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)  
 DE CARTILAGE INTERMEDIATE LAYER PROTEIN.  
 GN CILP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-ARTICULAR CARTILAGE;  
 RX MEDLINE; 98389785.  
 RA LORENZO P., NEAME P., SOMMARIN Y., HEINEGARD D.;  
 RT "Cloning and deduced amino acid sequence of a novel cartilage protein  
 RT (CILP) identifies a proform including a nucleotide  
 RL J. Biol. Chem. 273:23469-23475(1998).  
 DR EMBL; AF035408; AAC33838.1; -;  
 DR PFAM; PF00047; ig; 1;  
 DR PFAM; PF00090; tsp\_1; 1;  
 SQ SEQUENCE 1184 AA; 132538 MW; 3E6F9774 CRC32;

Query Match 65.4%; Score 34; DB 4; Length 1184;  
 Best Local Similarity 71.4%; Pred. No. 19e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DPTGHSY 9  
 DB 1068 DPLGHNY 1074

RESULT 15  
 O75862 PRELIMINARY; PRT; 347 AA.  
 AC O75862;  
 DT 01-NOV-1998 (TREMELrel. 08, Created)  
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)  
 DE MAGE-B1.  
 GN MAGE-B1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MUZYNY D., ARONSON A.D., ADAMS C., BRUNDAGE E., BUNAC C., CARVELLI K.,  
 RA CHACKO J., CHEN J., DI W., DING Y., DUGAN S., DURBIN J., FORCUM J.,  
 RA GANESH R., GARCIA C., GOODMAN M., GORRELL J.H., HAYWOOD M.,  
 RA HERNANDEZ J., JACKSON L., JIN S., KAMPAL R., KARPATY S., KOVAR C.,  
 RA LEAL B., LI Y., LICHTARGE O., LIU W., LOGAN O., LU J., LY T.,

RA MARTINEZ C., OSWAL G., PEREZ L., RASHID N.D., ROWLAND K., SAVAGE L.,  
RA SCHERER S.E., SHEN H., SIMON M., STOVALL K., TIMMS K.M., TODD J.,  
RA VO O., WILLIAMSON A., WORLEY K.C., YU W., CHINAULT C., NELSON D.,  
RA GIBBS R.A.;  
RT "Direct Submission";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC005185; AADI0634.1; -;  
DR PFAM; PF01454; MAGE; 1.  
SQ SEQUENCE 347 AA; 39037 MW; 48A34904 CRC32;

Query Match 65.4%; Score 34; DB 4; Length 347;  
Best Local Similarity 55.6%; Pred. No. 52;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EADPTGHSY 9  
| : : : : :  
Db 167 EDNPSGTY 175

Search completed: December 11, 1999, 01:16:39  
Job time: 3819 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 1999, 02:36:51 ; Search time 2290.15 Seconds  
(without alignments)  
7879.429 Million cell updates/sec

Title: US-08-819-669D-8  
Perfect score: 5674  
Sequence: 1 CCCGGGACCACTGGCATC.....TAATGATCTGGTGGATCC 5674

Scoring table: IDENTITY\_NUC

Searched: 679419 seqs, 1590154680 residues

Database :

GenEmbl:

1: gb\_ba1:  
2: gb\_ba2:  
3: gb\_om:  
4: gb\_ov:  
5: gb\_pat:  
6: gb\_ph:  
7: gb\_p11:  
8: gb\_p12:  
9: gb\_p13:  
10: gb\_p14:  
11: gb\_p15:  
12: gb\_p16:  
13: gb\_p17:  
14: gb\_p18:  
15: gb\_p19:  
16: gb\_p20:  
17: gb\_p21:  
18: gb\_p22:  
19: gb\_p23:  
20: gb\_p24:  
21: gb\_p25:  
22: gb\_p26:  
23: gb\_p27:  
24: gb\_p28:  
25: gb\_p29:  
26: gb\_p30:  
27: gb\_p31:  
28: gb\_p32:  
29: gb\_p33:  
30: gb\_p34:  
31: gb\_p35:  
32: gb\_p36:  
33: gb\_p37:  
34: gb\_p38:  
35: gb\_p39:  
36: gb\_p40:  
37: gb\_p41:  
38: gb\_p42:  
39: gb\_p43:  
40: gb\_p44:  
41: gb\_p45:  
42: gb\_p46:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5674	100.0	5674	5	I24013 Sequence 1

RESULT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45																																																																																																																																																								
LOCUS	I24013	100.0	5674	5	I36923	10	HSU82672	42	AF134576	9	HSU10687	11	U82696	9	HSU10688	9	HSU10690	9	HSU10689	9	I36922	5	AR007331	5	HUMAG1A	10	HSAF002994	11	AF030261	11	AF030261	10	HSAF002996	10	HSAF002996	9	HUMAGE2X	5	I36924	9	HUMAGE12X	20	1762.6	31.1	4204	9	HSU03735	21	1678.8	29.6	2531	5	I36928	22	1678.8	29.6	2531	5	I36929	23	1677	29.6	3871	9	HSU10691	24	1556.6	27.4	4327	10	HSU69568	25	1536.4	27.1	3680	9	HSU10692	26	1504.6	26.5	2305	5	I36932	27	1410.8	24.9	2226	5	I36931	28	1353.8	23.9	73360	11	HSU66083	29	1247.4	22.0	3839	9	HSU10693	30	1226.6	21.6	2931	9	HSU10694	31	1051	18.5	3672	9	HSU10686	32	988.8	17.4	1640	5	I36935	33	900.2	15.9	1810	5	I36935	34	866	15.3	3510	9	HSU10685	35	830.6	14.6	1947	5	I36934	36	815.2	14.4	1412	5	I36936	37	760	13.4	1022	9	HUMAGEA	38	756.8	13.3	1022	9	HUMAGEC	39	750.8	13.2	1068	5	I36930	40	733.6	12.9	1111	14	G06157	41	701.6	12.4	954	9	HSU10340	42	695.8	12.3	1019	9	HUMAGEB	43	664.2	11.7	76133	10	HSU71148	44	629.8	11.1	945	9	HSU10339	45	583.4	10.3	1107	5	I36938

## ALIGNMENTS

RESULT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45																																																																																																																																																								
LOCUS	I24013	100.0	5674	5	I36923	10	HSU82672	42	AF134576	9	HSU10687	11	U82696	9	HSU10688	9	HSU10690	9	HSU10689	9	I36922	5	AR007331	5	HUMAG1A	10	HSAF002994	11	AF030261	11	AF030261	10	HSAF002996	10	HSAF002996	9	HUMAGE2X	5	I36924	9	HUMAGE12X	20	1762.6	31.1	4204	9	HSU03735	21	1678.8	29.6	2531	5	I36928	22	1678.8	29.6	2531	5	I36929	23	1677	29.6	3871	9	HSU10691	24	1556.6	27.4	4327	10	HSU69568	25	1536.4	27.1	3680	9	HSU10692	26	1504.6	26.5	2305	5	I36932	27	1410.8	24.9	2226	5	I36931	28	1353.8	23.9	73360	11	HSU66083	29	1247.4	22.0	3839	9	HSU10693	30	1226.6	21.6	2931	9	HSU10694	31	1051	18.5	3672	9	HSU10686	32	988.8	17.4	1640	5	I36935	33	900.2	15.9	1810	5	I36935	34	866	15.3	3510	9	HSU10685	35	830.6	14.6	1947	5	I36934	36	815.2	14.4	1412	5	I36936	37	760	13.4	1022	9	HUMAGEA	38	756.8	13.3	1022	9	HUMAGEC	39	750.8	13.2	1068	5	I36930	40	733.6	12.9	1111	14	G06157	41	701.6	12.4	954	9	HSU10340	42	695.8	12.3	1019	9	HUMAGEB	43	664.2	11.7	76133	10	HSU71148	44	629.8	11.1	945	9	HSU10339	45	583.4	10.3	1107	5	I36938

LOCUS	I24013	Sequence 1 from patent US 5541104.	DNA	PAT	21-NOV-1996
DEFINITION	Sequence 1 from patent US 5541104.				
ACCESSION	I24013				
NID	G1603883				
VERSION	I24013.1	GI:1603883			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 5674)				
AUTHORS	Chen, Y., Stockert, E., Chen, Y., Garin-Chesa, P., Rettig, W.J., van der Bruggen, P., Boon-Falleur, T. and Old, L.J.				
TITLE	Monoclonal antibodies which bind to tumor rejection antigen precursor mage-1				
JOURNAL	Patent: US 5541104-A 1 30-JUL-1996;				
FEATURES	Location/Qualifiers				
Source	1 .5674				
BASE COUNT	1276 a, 1644 c 1569 g 1185 t				
ORIGIN					

Query Match 100.0%; Score 5674; DB 5; Length 5674;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5674; Conservative 0; Mismatches 0; Indels 0;

Qy	1	CCCGGGGACCACTGCGATCCCTCCCTCCCTTACACACCCCAATCCCTCCCTTTACGCCACCC	60
Db	1	CCCGGGGACCACTGCGATCCCTCCCTCCCTTACACACCCCAATCCCTCCCTTTACGCCACCC	60
Qy	61	ATCCAACATCTTCACGCTACCCCGACCCCAAGCCAGCCAGATCCGGTTCCACCCCTG	120
Db	61	ATCCAACATCTTCACGCTACCCCGACCCCAAGCCAGCCAGATCCGGTTCCACCCCTG	120
Qy	121	CTCTAACCCAGGGAAGCCAGGTGCCAGATGTGACGCCACTGACTTGAGCATTTAGTGG	180
Db	121	CTCTAACCCAGGGAAGCCAGGTGCCAGATGTGACGCCACTGACTTGAGCATTTAGTGG	180
Qy	181	TTAGAGAGAAGCGAGTTTTCGGTCTGAGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG	240
Db	181	TTAGAGAGAAGCGAGTTTTCGGTCTGAGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG	240
Qy	241	CCCAGCTCTGTAAGGAGCAAGGTGACATGCTGAGGGGAGGACTGAGAGCCACATTTACCCC	300
Db	241	CCCAGCTCTGTAAGGAGCAAGGTGACATGCTGAGGGGAGGACTGAGAGCCACATTTACCCC	300
Qy	301	AGATAGAGACCCCAATTAATCCCTTCATGCGCAGTCTTGAGACCATCTGGTGGTGGAGATTC	360
Db	301	AGATAGAGACCCCAATTAATCCCTTCATGCGCAGTCTTGAGACCATCTGGTGGTGGAGATTC	360
Qy	361	TCAGGCTGGGCCACCCACAGCCCCCTGCTGCTTAACCACTGGGGAATCGAAGTCAGAG	420
Db	361	TCAGGCTGGGCCACCCACAGCCCCCTGCTGCTTAACCACTGGGGAATCGAAGTCAGAG	420
Qy	421	CTCCGTGTGATCAGGGAAGGGTGTGTTAGGAGAGGGCAGCGTCCAGGGCTCTGCCAGACAT	480
Db	421	CTCCGTGTGATCAGGGAAGGGTGTGTTAGGAGAGGGCAGCGTCCAGGGCTCTGCCAGACAT	480
Qy	481	CATGCTCAGGATCTCAAGAGGGCTGAGGGTCCCTAGACCCCACTCCCGTGACCCCAAC	540
Db	481	CATGCTCAGGATCTCAAGAGGGCTGAGGGTCCCTAGACCCCACTCCCGTGACCCCAAC	540
Qy	541	CCCACTCCAAATGCTCACTCCGTCGACCAACCCCTCTTCATTGTTCATTCCACCCCCA	600
Db	541	CCCACTCCAAATGCTCACTCCGTCGACCAACCCCTCTTCATTGTTCATTCCACCCCCA	600
Qy	601	CCCCACATCCCCACCCCAATCCCTCAACCTGATGCCCATCCGCCCAGCCATTCCACCCCT	660
Db	601	CCCCACATCCCCACCCCAATCCCTCAACCTGATGCCCATCCGCCCAGCCATTCCACCCCT	660
Qy	661	CACCCCAACCCCAACCCCAACCCCACTCCACCCCAACCCCAAGCAGGATCCGGTTCCCG	720
Db	661	CACCCCAACCCCAACCCCAACCCCACTCCACCCCAACCCCAAGCAGGATCCGGTTCCCG	720
Qy	721	CCAGGAAACATCCGGTGCCCGGATGTGACGCCACTGACTTCGCATTTGGGGCAGAGA	780
Db	721	CCAGGAAACATCCGGTGCCCGGATGTGACGCCACTGACTTCGCATTTGGGGCAGAGA	780
Qy	781	GAGCAGAGTTTCCATTCTGAGGAGCGGGTAGATTTCGGCCGAAGAACTGACCCAGG	840
Db	781	GAGCAGAGTTTCCATTCTGAGGAGCGGGTAGATTTCGGCCGAAGAACTGACCCAGG	840
Qy	841	CTCTGTGAGGAGCAAGGTGAGAGCTGAGGAGGACTGAGGACCCCGCCACTCCAATA	900
Db	841	CTCTGTGAGGAGCAAGGTGAGAGCTGAGGAGGACTGAGGACCCCGCCACTCCAATA	900
Qy	901	GAGAGCCCCAAATATTCCAGCCCCCGCTTGTGCCAGCCCTTGCCCCCAGCGGGGAAGA	960
Db	901	GAGAGCCCCAAATATTCCAGCCCCCGCTTGTGCCAGCCCTTGCCCCCAGCGGGGAAGA	960
Qy	961	CGTCTCAGCCTGGGCTGCCCGCCAGACCCCTGCTCCAAAGGCTTTGAGAGACACCAAGTTC	1020
Db	961	CGTCTCAGCCTGGGCTGCCCGCCAGACCCCTGCTCCAAAGGCTTTGAGAGACACCAAGTTC	1020
Qy	1021	TTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAAGGCGAGGACTGGTTAGGAGAGG	1080



Db	2101	ACTGAGGCTGCCACTTCTTGGCCTCAAGAAATCAGAACGATGAGGGGACTCAGATTCGATGGGG	2160
Qy	2161	GTGGGACCCAGGCGCTCAAGAGGCTTACGCGGAGGAAGAGGAGGAGGACTCAGGGGACCTT	2220
Db	2161	GTGGGACCCAGGCGCTCAAGGCTTACGCGGAGGAAGAGGAGGAGGACTCAGGGGACCTT	2220
Qy	2221	GGATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCAGAGGTCACGGGACCGTGGCCACATATG	2280
Db	2221	GGATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCAGAGGTCACGGGACCGTGGCCACATATG	2280
Qy	2281	GCCCATATTTCCTGCATCTTTGAGTGCACAGACAGAGCTGTGTCTTGAGAGTGGGGCC	2340
Db	2281	GCCCATATTTCCTGCATCTTTGAGTGCACAGACAGAGCTGTGTCTTGAGAGTGGGGCC	2340
Qy	2341	TCAGGTCACACAGAGGAGGAGTTCCAGGATCCATATGCCCCAAGATGTGCCCCCTTCATG	2400
Db	2341	TCAGGTCACACAGAGGAGGAGTTCCAGGATCCATATGCCCCAAGATGTGCCCCCTTCATG	2400
Qy	2401	AGGACTGGGGATATCCCCGGCTCAGAAAGAGGAGTCCACACAGTCTGGCTGTGCCCTT	2460
Db	2401	AGGACTGGGGATATCCCCGGCTCAGAAAGAGGAGTCCACACAGTCTGGCTGTGCCCTT	2460
Qy	2461	TTAGTAGCTCTAGGGGACACAGATCAGGGATGGCGGTATGTTTCATTCTCACTTGTATACCA	2520
Db	2461	TTAGTAGCTCTAGGGGACACAGATCAGGGATGGCGGTATGTTTCATTCTCACTTGTATACCA	2520
Qy	2521	CAGCAGGAAGTTGGGGGCCCTCAGGAGATGGGGTCTTGGGTGAAGGGGGATGTCT	2580
Db	2521	CAGCAGGAAGTTGGGGGCCCTCAGGAGATGGGGTCTTGGGTGAAGGGGGATGTCT	2580
Qy	2581	ACTCATGTCAAGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT	2640
Db	2581	ACTCATGTCAAGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT	2640
Qy	2641	GAGACAGACAAGCTATTGGAATCCACACCCCGAACAAAGGGGTGAGCCCTGGACACC	2700
Db	2641	GAGACAGACAAGCTATTGGAATCCACACCCCGAACAAAGGGGTGAGCCCTGGACACC	2700
Qy	2701	TCACCCAGGATGTGGCTCTCTTTTCACCTCTCTTCCAGATCTCGGGCAGGTGAGACCT	2760
Db	2701	TCACCCAGGATGTGGCTCTCTTTTCACCTCTCTTCCAGATCTCGGGCAGGTGAGACCT	2760
Qy	2761	CATTCTCAGAGGTTGACTCAGGTCAACGTAGGACCCCCCATCTCGTCTAAAGACAGAGCG	2820
Db	2761	CATTCTCAGAGGTTGACTCAGGTCAACGTAGGACCCCCCATCTCGTCTAAAGACAGAGCG	2820
Qy	2821	GTCCCAGGATCTGCCATGCTTGGGTGAGGAACATGAGGAGAGACTGAGGGTACCCCGAG	2880
Db	2821	GTCCCAGGATCTGCCATGCTTGGGTGAGGAACATGAGGAGAGACTGAGGGTACCCCGAG	2880
Qy	2881	GACCAGAACACTCAGGAGACTGCACAGAAATCAGCCCTGCCCTGTGTCAACCCAGAG	2940
Db	2881	GACCAGAACACTCAGGAGACTGCACAGAAATCAGCCCTGCCCTGTGTCAACCCAGAG	2940
Qy	2941	AGCATGGGCTGGCCGCTCTGCCAGGTCTTCCGTTATCTCTGGGATCATTCATGTACGGG	3000
Db	2941	AGCATGGGCTGGCCGCTCTGCCAGGTCTTCCGTTATCTCTGGGATCATTCATGTACGGG	3000
Qy	3001	ACGGGAGGCCCTTGGTCTCAGAAAGCTGGCTCAGGTGAGGAGAGCGTCCAGGCC	3060
Db	3001	ACGGGAGGCCCTTGGTCTCAGAAAGCTGGCTCAGGTGAGGAGAGCGTCCAGGCC	3060
Qy	3061	CTGCCAGGAGTCAAGGTGAGGACCAAGCGGCACCTCACCCAGACACATTAATTCCAAT	3120
Db	3061	CTGCCAGGAGTCAAGGTGAGGACCAAGCGGCACCTCACCCAGGACACATTAATTCCAAT	3120
Qy	3121	GAATTTTGATATCTTGTCTGCCCTTCCCCAGGACCTTAGCAGCGTGTGCCAGATGTTT	3180
Db	3121	GAATTTTGATATCTTGTCTGCCCTTCCCCAGGACCTTAGCAGCGTGTGCCAGATGTTT	3180
Qy	3181	GTCCCTCTCTGCTTCCATTCCTTATCATGAGATGTGAATCTTGTATTTGGATTTCTCAG	3240
Db	3181	GTCCCTCTCTGCTTCCATTCCTTATCATGAGATGTGAATCTTGTATTTGGATTTCTCAG	3240

QY	3241	ACCAGCAAAAGGGCAGGATCCAGGCGCTGCCAGGAAAAATATAAGGGCCCTCGCGTGAGAA	3300
DB	3241	ACCAGCAAAAGGGCAGGATCCAGGCGCTGCCAGGAAAAATATAAGGGCCCTCGCGTGAGAA	3300
QY	3301	CAGAGGGGTCATCCACTGCATGAGAGTGGGATGTACAGAGTCCAGGCCACCCCTCTG	3360
DB	3301	CAGAGGGGTCATCCACTGCATGAGAGTGGGATGTACAGAGTCCAGGCCACCCCTCTG	3360
QY	3361	GTACACTGAGAAAGCCAGGGCTGTCTGCGGTCTGCACCTGTGAGGGCCGTGATTCCT	3420
DB	3361	GTACACTGAGAAAGCCAGGGCTGTCTGCGGTCTGCACCTGTGAGGGCCGTGATTCCT	3420
QY	3421	CTTCTGGAGCTCCAGGAACAGGCAAGTGGCCCTGTGCTGAGACAGTATCCTCAGGTC	3480
DB	3421	CTTCTGGAGCTCCAGGAACAGGCAAGTGGCCCTGTGCTGAGACAGTATCCTCAGGTC	3480
QY	3481	ACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACAA	3540
DB	3481	ACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACAA	3540
QY	3541	GGGCCCCACTGCCACAGGACATAGGACTCCACAGAGTCTGGCTCACCTCCCTACTG	3600
DB	3541	GGGCCCCACTGCCACAGGACATAGGACTCCACAGAGTCTGGCTCACCTCCCTACTG	3600
QY	3601	TCAGTCTGTAGANTCGACCTCTGCTGGCGGCTGTACCTGAGTACCTCTCACTTCCT	3660
DB	3601	TCAGTCTGTAGANTCGACCTCTGCTGGCGGCTGTACCTGAGTACCTCTCACTTCCT	3660
QY	3661	CCTTCAGGTTTCAGGGGACAGGCCAACCCAGGACAGGATTCCTGGAGGCCACAGAG	3720
DB	3661	CCTTCAGGTTTCAGGGGACAGGCCAACCCAGGACAGGATTCCTGGAGGCCACAGAG	3720
QY	3721	GAGCACCAGGAGAGATCTGTAAGTAGGCTTTGTTAGAGTCTCCAGAGTTCAAGTCTC	3780
DB	3721	GAGCACCAGGAGAGATCTGTAAGTAGGCTTTGTTAGAGTCTCCAGAGTTCAAGTCTC	3780
QY	3781	AGCTGAGGCTCTCACACACTCCCTCTCTCCCGAGGCTGTGGGTCTTCATTGCCACGCT	3840
DB	3781	AGCTGAGGCTCTCACACACTCCCTCTCTCCCGAGGCTGTGGGTCTTCATTGCCACGCT	3840
QY	3841	CCTGCCACACTCTCTGCTGCTGCCCTGACGAGATCATATGTCCTTTGAGCAGAGAG	3900
DB	3841	CCTGCCACACTCTCTGCTGCTGCCCTGACGAGATCATATGTCCTTTGAGCAGAGAG	3900
QY	3901	TCTGACTGCAAGCCTGAGGAAGCCTTCAGGCCCAACAGAGGCCCTGGCCCTGGTGTG	3960
DB	3901	TCTGACTGCAAGCCTGAGGAAGCCTTCAGGCCCAACAGAGGCCCTGGCCCTGGTGTG	3960
QY	3961	TGTGAGGCTGCCACCTCTCTCTCTCTGCTGCTGGGACCTCTGGAGAGGTGCC	4020
DB	3961	TGTGAGGCTGCCACCTCTCTCTCTCTGCTGCTGGGACCTCTGGAGAGGTGCC	4020
QY	4021	CACCTGTGGGTCAACAGATCCTCCCGAGAGTCTCTCAGGAGCCTCCGGCTTTCCCATAC	4080
DB	4021	CACCTGTGGGTCAACAGATCCTCCCGAGAGTCTCTCAGGAGCCTCCGGCTTTCCCATAC	4080
QY	4081	CATCACTTCACTCGACAGAGCAACCCAGTTCAGGGTTCACAGGCCCTGGAAGGAGGG	4140
DB	4081	CATCACTTCACTCGACAGAGCAACCCAGTTCAGGGTTCACAGGCCCTGGAAGGAGGG	4140
QY	4141	GCCAAAGCACTCTGTATCTGTGGAGTCTTGTTCGAGCAGTAACTACTAAGAAGGTGGC	4200
DB	4141	GCCAAAGCACTCTGTATCTGTGGAGTCTTGTTCGAGCAGTAACTACTAAGAAGGTGGC	4200
QY	4201	TGATTTGGTTGGTTTTCTGCTCCTCAAAATATCGAGCCAGGAGCCAGTACAAAAGGCAGA	4260
DB	4201	TGATTTGGTTGGTTTTCTGCTCCTCAAAATATCGAGCCAGGAGCCAGTACAAAAGGCAGA	4260
QY	4261	AATGCTGGAGAGTGTCAATCAAAATATCAAGCACTGTTTTCTGAGATCTTCGCAAGC	4320
DB	4261	AATGCTGGAGAGTGTCAATCAAAATATCAAGCACTGTTTTCTGAGATCTTCGCAAGC	4320

QY 4321 CTCAGTCTCTGAGTGGTCTTTTGGCAATGACGTGAAGAAAGCAGACACCCACCGGCCA 4380  
DB 4321 CTCAGTCTCTGAGTGGTCTTTTGGCAATGACGTGAAGAAAGCAGACACCCACCGGCCA 4380  
QY 4381 CTCCTATGTCTCTGACCTGCTAGCTCTCTCTATGATGGCTGCTGGGTGATATCA 4440  
DB 4381 CTCCTATGTCTCTGACCTGCTAGCTCTCTCTATGATGGCTGCTGGGTGATATCA 4440  
QY 4441 GATCATGCCCAAGACAGCTTCCTGATAATGTCTGCTGATGATGCAATGGAGGGCG 4500  
DB 4441 GATCATGCCCAAGACAGCTTCCTGATAATGTCTGCTGATGATGCAATGGAGGGCG 4500  
QY 4501 CCATGCTCTCTGAGGAGGAATCTGGGAGGAGCTGAGTGTGATGAGGTGATGAGGAG 4560  
DB 4501 CCATGCTCTCTGAGGAGGAATCTGGGAGGAGCTGAGTGTGATGAGGTGATGAGGAG 4560  
QY 4561 GGAGCAGAGTCCCTATGGGAGGCCAGGAGCTGCTCACCAAGATTTGGTGCAGGAAA 4620  
DB 4561 GGAGCAGAGTCCCTATGGGAGGCCAGGAGCTGCTCACCAAGATTTGGTGCAGGAAA 4620  
QY 4621 GTACCTGGAGTACGCGAGTCCGCGACAGTATCCCGACGCTATGAGTCTCTGTGGGT 4680  
DB 4621 GTACCTGGAGTACGCGAGTCCGCGACAGTATCCCGACGCTATGAGTCTCTGTGGGT 4680  
QY 4681 CCAAGGGCCCTCGCTGAACACAGCTATGTGAAGTCTCTGAGTGTGATCAAGGTCACT 4740  
DB 4681 CCAAGGGCCCTCGCTGAACACAGCTATGTGAAGTCTCTGAGTGTGATCAAGGTCACT 4740  
QY 4741 GCAAGAGTTCGCTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGGAGGAGAG 4800  
DB 4741 GCAAGAGTTCGCTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGGAGGAGAG 4800  
QY 4801 GGAGTCTGAGCATGAGTGTGACCAAGCCAGTGGGAGGAGTGGCCAGTGCACCTT 4860  
DB 4801 GGAGTCTGAGCATGAGTGTGACCAAGCCAGTGGGAGGAGTGGCCAGTGCACCTT 4860  
QY 4861 CCAGGGCCGCTCCAGCAGCTTCCCTGCTGCTGAGTGCACATGAGGCCATCTTCACTCT 4920  
DB 4861 CCAGGGCCGCTCCAGCAGCTTCCCTGCTGCTGAGTGCACATGAGGCCATCTTCACTCT 4920  
QY 4921 GAAAGAGCGGTGAGTGTCTCAGTAGTAGTCTTCTGCTTATTTGGGTGAGTGGAGATT 4980  
DB 4921 GAAAGAGCGGTGAGTGTCTCAGTAGTAGTCTTCTGCTTATTTGGGTGAGTGGAGATT 4980  
QY 4981 TATCTTTGTTCTCTTTTGGAAATCTGTAATGTTTCTTTTAAAGGATGTTGAATGAAC 5040  
DB 4981 TATCTTTGTTCTCTTTTGGAAATCTGTAATGTTTCTTTTAAAGGATGTTGAATGAAC 5040  
QY 5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTACACAGTCTCTGTTATATAGTTTAAAGG 5100  
DB 5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTACACAGTCTCTGTTATATAGTTTAAAGG 5100  
QY 5101 TAAAGTCTGTGTTTATTCAGATTTGGGAATCCATCTATTTTGTGAATTTGGGATAAT 5160  
DB 5101 TAAAGTCTGTGTTTATTCAGATTTGGGAATCCATCTATTTTGTGAATTTGGGATAAT 5160  
QY 5161 AACAGCAGTGAATTAAGTACTTGAATATGAAATGTAATAATGAGCAGTAAATAGATGAGATA 5220  
DB 5161 AACAGCAGTGAATTAAGTACTTGAATATGAAATGTAATAATGAGCAGTAAATAGATGAGATA 5220  
QY 5221 AGAATCTAAAGAAATTAAGAGATAGTCAATTTCTGCTTATACCTCAGTCTATTTCTGATA 5280  
DB 5221 AGAATCTAAAGAAATTAAGAGATAGTCAATTTCTGCTTATACCTCAGTCTATTTCTGATA 5280  
QY 5281 ATTTTAAAGATATATGATACATCTGATTTCTTGGCTTCTTTGAGATGTAAGAGAAAT 5340  
DB 5281 ATTTTAAAGATATATGATACATCTGATTTCTTGGCTTCTTTGAGATGTAAGAGAAAT 5340  
QY 5341 TAAATCTGAATAAAGAAATCTTCTGTTCTGCTGCTGCTCTTTTCTCTCCATGCTGAGCA 5400  
DB 5341 TAAATCTGAATAAAGAAATCTTCTGTTCTGCTGCTGCTCTTTTCTCTCCATGCTGAGCA 5400  
QY 5401 TCTGCTTTTGGAGGGCCCTGGGTTAGTAGTGGAGATGCTTAAGTAAGCCAGACTCATAC 5460

DB 5401 TCTGCTTTTGGAGGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATAC 5460  
QY 5461 CCACCCATAGGTCGTAGAGTCTAGGAGCTGACGTACAGTAATCGAGTGGCAAGATGTC 5520  
DB 5461 CCACCCATAGGTCGTAGAGTCTAGGAGCTGACGTACAGTAATCGAGTGGCAAGATGTC 5520  
QY 5521 CTCTAAAGATGTAGGAAAAAGTGAAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580  
DB 5521 CTCTAAAGATGTAGGAAAAAGTGAAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580  
QY 5581 GTGAGTGTCAATGCCCTGAGCTGGGCAATTTGGGCTTTTGGAAACTGCAGTTCCTTCT 5640  
DB 5581 GTGAGTGTCAATGCCCTGAGCTGGGCAATTTGGGCTTTTGGAAACTGCAGTTCCTTCT 5640  
QY 5641 GGGGAGCTGATTTGTAATGATCTTTGGGTGATCC 5674  
DB 5641 GGGGAGCTGATTTGTAATGATCTTTGGGTGATCC 5674

RESULT 2  
LOCUS 136923 5674 bp DNA PAT 21-APR-1997  
DEFINITION Sequence 8 from patent US 5612201.  
ACCESSION 136923  
NID 92084883  
VERSION 136923.1 GI:2084883  
KEYWORDS  
SOURCE Unknown:  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5674)  
AUTHORS De Plaen,E., Boon-Falleur,T., Lethe ,B., Szikora,J., De Smet,C. and Chomez,P.  
TITLE Isolated nucleic acid molecules useful in determining expression of a tumor rejection antigen precursor  
JOURNAL Patent: US 5612201-A 8 18-MAR-1997;  
FEATURES Location/Qualifiers  
source 1. 5674  
BASE COUNT 1276 a, 1644 c 1569 g 1185 t  
ORIGIN

Query Match 100.0%; Score 5674; DB 5; Length 5674;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGGGGACACACTGGCATCCCTCCCTACCCACCCCAATCCCTCCCTTTACGCCACCC 60  
DB 1 CCGGGGACACACTGGCATCCCTCCCTACCCACCCCAATCCCTCCCTTTACGCCACCC 60  
QY 61 ATCAAAACATCTTCACGCTCACCCCAAGCCAGCCAGGAGAGATCCGGTTCACCCCTG 120  
DB 61 ATCAAAACATCTTCACGCTCACCCCAAGCCAGCCAGGAGAGATCCGGTTCACCCCTG 120  
QY 121 CTCTCAACCCAGGAAGCCAGAGTCCCGATGTGACGCCACCTGAGATCGGTGGAGGAGGGG 180  
DB 121 CTCTCAACCCAGGAAGCCAGAGTCCCGATGTGACGCCACCTGAGATCGGTGGAGGAGGGG 180  
QY 181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGGGCTTTGAGATCGGTGGAGGAGGGGG 240  
DB 181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGGGCTTTGAGATCGGTGGAGGAGGGGG 240  
QY 241 CCCAGCTCTTAAAGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGAGCCACTTACCCC 300  
DB 241 CCCAGCTCTTAAAGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGAGCCACTTACCCC 300  
QY 301 AGATAGAGGACCCCAATAATCCCTTCATGCCAGTCTCGACCATCTGGTGGTGGACTTC 360  
DB 301 AGATAGAGGACCCCAATAATCCCTTCATGCCAGTCTCGACCATCTGGTGGTGGACTTC 360  
QY 361 TCAGGCTGGCCACCCCGCCAGCCCTTCTGCTTAAACCCACTGGGAGCTCGAAGTCAGAG 420

||||| 361 TCAGGCTGGGCGCCACCCAGCCCTTGGCTTTAAACCACTGGGGACTGGAATCAGAG 420  
QY 421 CTCGCTGTATCAGGGAAGGCTGCTTGGAGAGGGGAGGCTCCAGGCTCTGCCAGACAT 480  
Db 421 CTCGCTGTATCAGGGAAGGCTGCTTGGAGAGGGGAGGCTCCAGGCTCTGCCAGACAT 480  
QY 481 CATGCTCAGGATTCAGGAGGGCTGAGGCTCCCTAAGACCCCACTCCGGTGACCCAAAC 540  
Db 481 CATGCTCAGGATTCAGGAGGGCTGAGGCTCCCTAAGACCCCACTCCGGTGACCCAAAC 540  
QY 541 CCCACTCCAATGCTCACTCCCGTGACCCACCCCTTCTTCAATGTCTATTCACACCCCA 600  
Db 541 CCCACTCCAATGCTCACTCCCGTGACCCACCCCTTCTTCAATGTCTATTCACACCCCA 600  
QY 601 CCCACATCCCCACCCCACTCCCTCAACCTGTATGCCATCCCGCCAGCCATTCACCCCT 660  
Db 601 CCCACATCCCCACCCCACTCCCTCAACCTGTATGCCATCCCGCCAGCCATTCACCCCT 660  
QY 661 CACCCCAACATCCGGGTGCGCGGATGTAGCCCACTGACTTGGCCATTTGGGGCAGAGA 780  
Db 661 CACCCCAACATCCGGGTGCGCGGATGTAGCCCACTGACTTGGCCATTTGGGGCAGAGA 780  
QY 781 GAAGCGAGGTTTCCATTTGAGGACCGGCTAGAGTTTCGGCCGAAGAACTGACCCAGG 840  
Db 781 GAAGCGAGGTTTCCATTTGAGGACCGGCTAGAGTTTCGGCCGAAGAACTGACCCAGG 840  
QY 841 CTCTGTGAGGAGCAAGGTGAGAGGCTGAGGAGGACTGAGGACCCGCCACTCCAATA 900  
Db 841 CTCTGTGAGGAGCAAGGTGAGAGGCTGAGGAGGACTGAGGACCCGCCACTCCAATA 900  
QY 901 GAGAGCCCCAAATTCAGGCCCCGCTTGTGCTGCCAGCCCTGCGCCACCCGGGAGAG 960  
Db 901 GAGAGCCCCAAATTCAGGCCCCGCTTGTGCTGCCAGCCCTGCGCCACCCGGGAGAG 960  
QY 961 CGTCTAGGCTGGGCTGCCCGCCAGACCCCTGCTCCAAAGCCTTGAGAGACACCAAGTTC 1020  
Db 961 CGTCTAGGCTGGGCTGCCCGCCAGACCCCTGCTCCAAAGCCTTGAGAGACACCAAGTTC 1020  
QY 1021 TTCTCCCAAGCTCTGGAATCAGAGTTGCTGTGACAGGCGAGGCTGCTTGGAGAGG 1080  
Db 1021 TTCTCCCAAGCTCTGGAATCAGAGTTGCTGTGACAGGCGAGGCTGCTTGGAGAGG 1080  
QY 1081 GCAGGCGACAGGCTCTGCCAGGCATCAAGATCAGCACCCCAAGAGGGGCTGTGGGCC 1140  
Db 1081 GCAGGCGACAGGCTCTGCCAGGCATCAAGATCAGCACCCCAAGAGGGGCTGTGGGCC 1140  
QY 1141 CCAAGACTGCACCTCAATCCCACTCCACCCCACTCCGATTCGATTCCTCCACCCCAACC 1200  
Db 1141 CCAAGACTGCACCTCAATCCCACTCCACCCCACTCCGATTCGATTCCTCCACCCCAACC 1200  
QY 1201 CCCATCTCCTCAGCTACACCTTCCACCCCACTCCCTACTCTCTACTCCCTCAGCTGACC 1260  
Db 1201 CCCATCTCCTCAGCTACACCTTCCACCCCACTCCCTACTCTCTACTCCCTCAGCTGACC 1260  
QY 1261 ACCCTCAGGCCCCAGCACCCAGCCCACTTCTGCGACCTCACCCTCACTGCCCCCAAC 1320  
Db 1261 ACCCTCAGGCCCCAGCACCCAGCCCACTTCTGCGACCTCACCCTCACTGCCCCCAAC 1320  
QY 1321 CCCACCTCATCTCTCATGTGCCCCACCTCCCATGCCCCCTCCCTTCTGGCAGAGATCC 1380  
Db 1321 CCCACCTCATCTCTCATGTGCCCCACCTCCCATGCCCCCTCCCTTCTGGCAGAGATCC 1380  
QY 1381 GGTTCGCTCTCAACCCAGGAGGCTTGGTGGGCGGCTGTGAAACCACTGACT 1440  
Db 1381 GGTTCGCTCTCAACCCAGGAGGCTTGGTGGGCGGCTGTGAAACCACTGACT 1440  
QY 1441 TGAACCTCACAGATCTCAGAGAACCAAGTTTCAATTTATGGTTCTGAGGGGCGGCTGAG 1500  
|||||

Db 1441 TGAACCTCACAGATCTCAGAGAACCAAGTTTCAATTTATGGTTCTGAGGGGCGGCTTGA 1500  
QY 1501 ATCCACTAGAGGAGTGGTTTTAGGCTCTGTGAGAGGCAAGGTGAGATGCTGAGGAGG 1560  
Db 1501 ATCCACTAGGAGTGGTTTTAGGCTCTGTGAGAGGCAAGGTGAGATGCTGAGGAGG 1560  
QY 1561 ACTGAGAGGACACACACCCAGGTAGATGGCCCAAAATGATCCAGTACACCCCTGCTG 1620  
Db 1561 ACTGAGAGGACACACACCCAGGTAGATGGCCCAAAATGATCCAGTACACCCCTGCTG 1620  
QY 1621 CCAGCCTTGACACACCCGCGCAGGACAGATGCTCAGCTGGACACACCCCGTCCCGTCC 1680  
Db 1621 CCAGCCTTGACACACCCGCGCAGGACAGATGCTCAGCTGGACACACCCCGTCCCGTCC 1680  
QY 1681 CACTGCCACTTAAACCCACAGGCAATCTGTAGTTCATAGCTTATGTAGCCGGGAGGTT 1740  
Db 1681 CACTGCCACTTAAACCCACAGGCAATCTGTAGTTCATAGCTTATGTAGCCGGGAGGTT 1740  
QY 1741 GGTGAGGAGGAGGCGCCAGGCATCAAGTCCAGCATCCGCCGCAATTAGGTCAGG 1800  
Db 1741 GGTGAGGAGGAGGCGCCAGGCATCAAGTCCAGCATCCGCCGCAATTAGGTCAGG 1800  
QY 1801 ACCCTGGAGGAGTCTGAGGTTCCCCACCCACACCTGTCTCTCATCTCCACCCGACC 1860  
Db 1801 ACCCTGGAGGAGTCTGAGGTTCCCCACCCACACCTGTCTCTCATCTCCACCCGACC 1860  
QY 1861 CCATCTCATTTCCCATACCTACCCCTTACCCCAACCTCATCTTGTGAGAACTCCCTGCTG 1920  
Db 1861 CCATCTCATTTCCCATACCTACCCCTTACCCCAACCTCATCTTGTGAGAACTCCCTGCTG 1920  
QY 1921 TCAACCCACGGAAGCCAGGGAATGCGGCGCAGGCACTCGGATTTGACGTCCCATCCA 1980  
Db 1921 TCAACCCACGGAAGCCAGGGAATGCGGCGCAGGCACTCGGATTTGACGTCCCATCCA 1980  
QY 1981 GGGTCTGATGAGGGAAGGGCTTGAACAGGGCTCAGGAGGAGGAGGAGGAGGAGGAGG 2040  
Db 1981 GGGTCTGATGAGGGAAGGGCTTGAACAGGGCTCAGGAGGAGGAGGAGGAGGAGGAGG 2040  
QY 2041 TGCAGATGAGGAGGCTCAGAGGACCCAGCACCTTAGGACACCCGACCCCTGCTGTAG 2100  
Db 2041 TGCAGATGAGGAGGCTCAGAGGACCCAGCACCTTAGGACACCCGACCCCTGCTGTAG 2100  
QY 2101 ACTGAGGCTGCCACTTCTGGCTCAAGAACTCAGAACGATGGGAGTCTGAGATGGGG 2160  
Db 2101 ACTGAGGCTGCCACTTCTGGCTCAAGAACTCAGAACGATGGGAGTCTGAGATGGGG 2160  
QY 2161 GTGGACCCAGGCTGCAAGGCTTACGCGAGGAGGAGGAGGAGGAGTCTCAGGAGGACCTT 2220  
Db 2161 GTGGACCCAGGCTGCAAGGCTTACGCGAGGAGGAGGAGGAGGAGTCTCAGGAGGACCTT 2220  
QY 2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGCGACGTCGCCACATATG 2280  
Db 2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGCGACGTCGCCACATATG 2280  
QY 2281 GCCCATATTTCTGCTCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAAGTGGGCC 2340  
Db 2281 GCCCATATTTCTGCTCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAAGTGGGCC 2340  
QY 2341 TCAGGTCAACAGAGGAGGAGTTCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATG 2400  
Db 2341 TCAGGTCAACAGAGGAGGAGTTCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATG 2400  
QY 2401 AGGACTGGGATATCCCGGCTCAGAAAGAGGACCTCCACACAGTCTGGCTGTGCCCTT 2460  
Db 2401 AGGACTGGGATATCCCGGCTCAGAAAGAGGACCTCCACACAGTCTGGCTGTGCCCTT 2460  
QY 2461 TTAGTAGCTCTAGGGGACAGATCAGGATGGCGGTATGTTCCATTTCTCACTTTGTACCA 2520  
Db 2461 TTAGTAGCTCTAGGGGACAGATCAGGATGGCGGTATGTTCCATTTCTCACTTTGTACCA 2520  
QY 2521 CAGCAGGAAGTTGGGGGGCCCTCAGGAGATGGGGTCTTGGGGTAAAGGGGGATGTCT 2580  
Db 2521 CAGCAGGAAGTTGGGGGGCCCTCAGGAGATGGGGTCTTGGGGTAAAGGGGGATGTCT 2580  
|||||

QY 2581 ACTGATGTCAGGGAATTTGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT 2640  
DB 2581 ACTGATGTCAGGGAATTTGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT 2640  
QY 2641 GAGACAGACAAGGCTATTGGAATTCACACCCCGAGAACCAAGGGGTGAGCCCTGGACACC 2700  
DB 2641 GAGACAGACAAGGCTATTGGAATTCACACCCCGAGAACCAAGGGGTGAGCCCTGGACACC 2700  
QY 2701 TCACCCAGGATGTCGCTTCTTTTTCACCTCTCTTCCAGATCTGGGCGAGGTGAGGACCT 2760  
DB 2701 TCACCCAGGATGTCGCTTCTTTTTCACCTCTCTTCCAGATCTGGGCGAGGTGAGGACCT 2760  
QY 2761 CATTTCTCAGAGGCTGACTCAGGTCAACGTAGGAGCCCGCCCATCTGCTCTAAAGACAGAGCG 2820  
DB 2761 CATTTCTCAGAGGCTGACTCAGGTCAACGTAGGAGCCCGCCCATCTGCTCTAAAGACAGAGCG 2820  
QY 2821 GTCCAGGATCTGCCATGCTGCGGTGAGGAACATGAGGAGGACTGAGGATACCCAG 2880  
DB 2821 GTCCAGGATCTGCCATGCTGCGGTGAGGAACATGAGGAGGACTGAGGATACCCAG 2880  
QY 2881 GACCAAGACACTGAGGAGACTGCACAGAAATCAGCCCTGCCCTGCTGTACCCACAGAG 2940  
DB 2881 GACCAAGACACTGAGGAGACTGCACAGAAATCAGCCCTGCCCTGCTGTACCCACAGAG 2940  
QY 2941 AGCATGGGCTGGGCGCTCTGCCGAGGTCTTCCGTTATCTCTGGGATCAATTGATGTCAGGG 3000  
DB 2941 AGCATGGGCTGGGCGCTCTGCCGAGGTCTTCCGTTATCTCTGGGATCAATTGATGTCAGGG 3000  
QY 3001 AGGGGAGGCTTGGTCTGAGAAGGCTGGCTCAGGTCAAGTAGAGGAGCGTCCAGGCG 3060  
DB 3001 AGGGGAGGCTTGGTCTGAGAAGGCTGGCTCAGGTCAAGTAGAGGAGCGTCCAGGCG 3060  
QY 3061 CTGCGAGGAGTCAAGGTGAGGAGCAAGCGGCGACCTCACCCAGGACACATTAATTCACAT 3120  
DB 3061 CTGCGAGGAGTCAAGGTGAGGAGCAAGCGGCGACCTCACCCAGGACACATTAATTCACAT 3120  
QY 3121 GAATTTGATATCTCTGCTGCTTCCCAAGGACCTAGGACGCTGTGGCCAGATGTTT 3180  
DB 3121 GAATTTGATATCTCTGCTGCTTCCCAAGGACCTAGGACGCTGTGGCCAGATGTTT 3180  
QY 3181 GTCCCTCTGCTTCCATCTTATCATGATGTGAATCTTGAATTTGGATTTCTCAG 3240  
DB 3181 GTCCCTCTGCTTCCATCTTATCATGATGTGAATCTTGAATTTGGATTTCTCAG 3240  
QY 3241 ACCAGCAAAAGGCGAGGATCCAGGCGCTCCAGGAAAAATATAAGGGCCCTCGCTGAGAA 3300  
DB 3241 ACCAGCAAAAGGCGAGGATCCAGGCGCTCCAGGAAAAATATAAGGGCCCTCGCTGAGAA 3300  
QY 3301 CAGAGGGGTATCTCACTGATGAGAGTGGGATGTACAGAGTCCAGCCACCCCTCCTG 3360  
DB 3301 CAGAGGGGTATCTCACTGATGAGAGTGGGATGTACAGAGTCCAGCCACCCCTCCTG 3360  
QY 3361 GTAGCACTGAGAAGCCAGGGCTGTGCTTCCGCTGTGACCCCTGAGGCGCCGTGGATTCT 3420  
DB 3361 GTAGCACTGAGAAGCCAGGGCTGTGCTTCCGCTGTGACCCCTGAGGCGCCGTGGATTCT 3420  
QY 3421 CTTCTGAGCTCCAGGAACAGGAGTCCAGGCTTGGTCTGAGACAGTATCTCAGGTC 3480  
DB 3421 CTTCTGAGCTCCAGGAACAGGAGTCCAGGCTTGGTCTGAGACAGTATCTCAGGTC 3480  
QY 3481 ACAGAGCAGAGGATGCACAGGCTGTCAGCAGTGAATGTTTTGCCCTGAATGCACACCAA 3540  
DB 3481 ACAGAGCAGAGGATGCACAGGCTGTCAGCAGTGAATGTTTTGCCCTGAATGCACACCAA 3540  
QY 3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGGCTCACCTCCCTACTG 3600  
DB 3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGGCTCACCTCCCTACTG 3600  
QY 3601 TCAGTCTCTAGAAATCGACCTCTGCTGGCGGCTGTACCTCAGTACCTCTCACTTCT 3660  
DB 3601 TCAGTCTCTAGAAATCGACCTCTGCTGGCGGCTGTACCTCAGTACCTCTCACTTCT 3660

QY 3661 CTTTCAGGTTTTCAGGGACAGGCCAACCCAGAGACAGGATTCCTCTGAGGCCACAGAG 3720  
DB 3661 CTTTCAGGTTTTCAGGGACAGGCCAACCCAGAGACAGGATTCCTCTGAGGCCACAGAG 3720  
QY 3721 GAGCACCAGGAGAGATCTGTAAGTAGCCCTTTGTAGAGTCTCAAGGTTTCAGTTCTC 3780  
DB 3721 GAGCACCAGGAGAGATCTGTAAGTAGCCCTTTGTAGAGTCTCAAGGTTTCAGTTCTC 3780  
QY 3781 AGCTCAGGCTCTCACACACTCCCTCTCTCCCGAGCCCTGTGGTCTTCATTGCCAGCT 3840  
DB 3781 AGCTCAGGCTCTCACACACTCCCTCTCTCCCGAGCCCTGTGGTCTTCATTGCCAGCT 3840  
QY 3841 CTTGCCACACTCTCTGCTGCTGCCCTGACGAGAGTCAATCATCTCTTTGAGCAGAGAG 3900  
DB 3841 CTTGCCACACTCTCTGCTGCTGCCCTGACGAGAGTCAATCATCTCTTTGAGCAGAGAG 3900  
QY 3901 TCTGCACTGCAAGCCTGAGGAAGCCCTTTGAGGCCCAACAGAGGCCCTGGGCTGGTGTG 3960  
DB 3901 TCTGCACTGCAAGCCTGAGGAAGCCCTTTGAGGCCCAACAGAGGCCCTGGGCTGGTGTG 3960  
QY 3961 TGTGCAAGCTGCACT 4020  
DB 3961 TGTGCAAGCTGCACT 4020  
QY 4021 CACTGCTGGTCAACAGATCT 4080  
DB 4021 CACTGCTGGTCAACAGATCT 4080  
QY 4081 CATCAACTTCACTCAGAGAGCAACCCAGTGGGTTTCCAGCAGCCGTTGAAGAGAGGG 4140  
DB 4081 CATCAACTTCACTCAGAGAGCAACCCAGTGGGTTTCCAGCAGCCGTTGAAGAGAGGG 4140  
QY 4141 GCCAAGCACT 4200  
DB 4141 GCCAAGCACT 4200  
QY 4201 TGATTTGGTGGTTTCT 4260  
DB 4201 TGATTTGGTGGTTTCT 4260  
QY 4261 AATGCTGAGAGTGTCAATCAAAATTAAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4320  
DB 4261 AATGCTGAGAGTGTCAATCAAAATTAAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4320  
QY 4321 CTCTGAGTCTCTGAGCTGGTCTTTGGCATTTGACCTGAAGAGCAGACCCACCGGCA 4380  
DB 4321 CTCTGAGTCTCTGAGCTGGTCTTTGGCATTTGACCTGAAGAGCAGACCCACCGGCA 4380  
QY 4381 CTCTGATGCTCTGATGCTGCT 4440  
DB 4381 CTCTGATGCTCTGATGCTGCT 4440  
QY 4441 GATCATGCTCAGAGCAGGCTTCT 4500  
DB 4441 GATCATGCTCAGAGCAGGCTTCT 4500  
QY 4501 CCATGCTCTCTGAGGAGAAATCTGGAGAGCTGAGTGTGATGAGGAGTGTATGATGGGAG 4560  
DB 4501 CCATGCTCTCTGAGGAGAAATCTGGAGAGCTGAGTGTGATGAGGAGTGTATGATGGGAG 4560  
QY 4561 GGAGCAGAGTGCCTATGGGAGCCCGAGGAAGTGTCTACCCCAAGATTTGGTGCAGGAAA 4620  
DB 4561 GGAGCAGAGTGCCTATGGGAGCCCGAGGAAGTGTCTACCCCAAGATTTGGTGCAGGAAA 4620  
QY 4621 GTACCTGAGTACGCGAGTGGCGGACAGTGTATCCCGACGCTATGAGTCTCTGTTGGGT 4680  
DB 4621 GTACCTGAGTACGCGAGTGGCGGACAGTGTATCCCGACGCTATGAGTCTCTGTTGGGT 4680  
QY 4681 CCAAGGGCCCTCTCTGAAACAGCTATGTGAAGTCTCTGATGTATGATCAAGGTCT 4740  
DB 4681 CCAAGGGCCCTCTCTGAAACAGCTATGTGAAGTCTCTGATGTATGATCAAGGTCT 4740  
QY 4741 GCAAGAGTTCGCTTTTCTTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4800

```
|||||
Db 4741 GCAGAGTTCGCTTTCTTCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGAG 4800
Qy 4801 GGAGTCTGAGCATGAGTTGACCAAGCCAGTGGAGGGGAGTGGCCAGTGCACCTT 4860
Db 4801 GGAGTCTGAGCATGAGTTGACCAAGCCAGTGGAGGGGAGTGGCCAGTGCACCTT 4860
Qy 4861 CCAGGGCGGCTCCAGCAGCTTCCCTGCTGCTGATGACATGAGGCCCATCTTTCACCT 4920
Db 4861 CCAGGGCGGCTCCAGCAGCTTCCCTGCTGCTGATGACATGAGGCCCATCTTTCACCT 4920
Qy 4921 GAAGAGAGCGTCACTGCTTCAGTAGTAGTTCCTGCTTCTTCTTCTGAGTGGAGATT 4980
Db 4921 GAAGAGAGCGTCACTGCTTCAGTAGTAGTTCCTGCTTCTTCTTCTGAGTGGAGATT 4980
Qy 4981 TATCTTTGCTCTTTTGGAAATGTTCAAAATGTTTTTTTAAAGGATGGTTGAATGAAC 5040
Db 4981 TATCTTTGCTCTTTTGGAAATGTTCAAAATGTTTTTTTAAAGGATGGTTGAATGAAC 5040
Qy 5041 TTCAGCATCCAAAGTTTATGAATACAGCAGTGCACACAGTTCCTGCTGATATAGTTAAGG 5100
Db 5041 TTCAGCATCCAAAGTTTATGAATACAGCAGTGCACACAGTTCCTGCTGATATAGTTAAGG 5100
Qy 5101 TAAGAGTCTGCTGTTTATTCAGATGGGAATCCATCTTATTTGGAATGGGATAAT 5160
Db 5101 TAAGAGTCTGCTGTTTATTCAGATGGGAATCCATCTTATTTGGAATGGGATAAT 5160
Qy 5161 AACAGCAGTGAATAAGTACCTAGAAATGTCAAAATGAGCAGTAAATAGATGAGATAA 5220
Db 5161 AACAGCAGTGAATAAGTACCTAGAAATGTCAAAATGAGCAGTAAATAGATGAGATAA 5220
Qy 5221 AGAACTAAAGAAATTAAGAGATAGTCAATCTTCCCTTATACCTCAGTCTATTTCTGATAA 5280
Db 5221 AGAACTAAAGAAATTAAGAGATAGTCAATCTTCCCTTATACCTCAGTCTATTTCTGATAA 5280
Qy 5281 ATTTTAAAGATATATGATACCTAGTATTCCTGCTGCTTCTTCTGAGATGTAAGAGAAAT 5340
Db 5281 ATTTTAAAGATATATGATACCTAGTATTCCTGCTGCTTCTTCTGAGATGTAAGAGAAAT 5340
Qy 5341 TAAATCTGAATAAGAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5400
Db 5341 TAAATCTGAATAAGAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5400
Qy 5401 TCTGCTTTTGGAAAGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATAC 5460
Db 5401 TCTGCTTTTGGAAAGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATAC 5460
Qy 5461 CCACCCATAGGCTGCTAGAGTCTAGGAGCTGAGTCACTGCTGCTGCTGCTGCTGCTGCTG 5520
Db 5461 CCACCCATAGGCTGCTAGAGTCTAGGAGCTGAGTCACTGCTGCTGCTGCTGCTGCTGCTG 5520
Qy 5521 CTCTAAGATCTAGGAAAGCTGAGAGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5580
Db 5521 CTCTAAGATCTAGGAAAGCTGAGAGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5580
Qy 5581 GTGAGTGTCAATCCCTGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5640
Db 5581 GTGAGTGTCAATCCCTGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5640
Qy 5641 GGGGAGCTGATGTAATGATCTTGGGTGGATCC 5674
Db 5641 GGGGAGCTGATGTAATGATCTTGGGTGGATCC 5674
```

```
RESULT 3
HSU82672/c
LOCUS HSU82672 156854 bp DNA PRI 13-MAY-1997
DEFINITION Human chromosome X clone Qc15B1, complete sequence.
ACCESSION U82672
NID 92078526
VERSION U82672.1 GI:2078526
KEYWORDS HTG.
SOURCE human.
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 156854)
AUTHORS Gloeckner, G., Rump, A., Nordstiek, G., Hinzmann, B., Kioschis, P.,
Haiss, N., Poustka, A., Bauer, D., Drescher, B., Knob, A. and
Rosenthal, A.
TITLE Sequencing and analysis of a region in Xq28 containing MAGE-1 and a
putative Zinc Finger Protein
JOURNAL Unpublished
AUTHORS Gernot, G.
REFERENCE 2 (bases 1 to 156854)
JOURNAL Direct Submission
AUTHORS Submitted (19-DEC-1996) Genome Analysis, IMB, Beutenbergstr. 11,
Jena 07745, Germany
FEATURES
Location/Qualifiers
source 1..156854
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xq28"
1..40357
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Qc15B1"
1057..1295
/repeat_region
12504..2666
/rpt_family="Alu-Sz"
2504..2666
/rpt_family="MIR"
complement(4262..5191)
/organism="Homo sapiens"
/product="MAGE-1 Protein"
complement(4262..5191)
/gene="MAGE1"
complement(4262..5191)
/gene="MAGE1"
/codon_start=1
/product="MAGE-1 Protein"
/protein_id="AAB54061.1"
/db_xref="PID:Q2078527"
/db_xref="GI:2078527"
/translation="MSLEQSLHCKPEEALEAOEALGLVCVQAAASSSPVLVLTLE
EVTAGSTDPQSQGASAPTTINFTQRPQEGSSREEGSTSCILLESFRAVI
TKVADLVGFLLLKRYAREPVTAKMLESVKYKHCFFIFGKASLSQLVFDIVK
EADPTGHSYVLVTCGLSYDGLDNOIMPKTGTFLIIVLVMIAMEGHAEETWEEL
SVMEVDGRHSAYGEPKLLTODLVQEKYLEVRQVDSDPARVEFLWGPALAEYSY
VKVLEYVYKVSARVRFPPSLREALREEEGV"
9623..9892
/rpt_family="Alu-Sz"
complement(10698..10978)
/rpt_family="Alu-Sx"
complement(12218..12345)
/rpt_family="MER42C"
complement(12636..12851)
/rpt_family="MER42A"
complement(12909..13060)
/rpt_family="MER42A"
complement(13455..13825)
/rpt_family="MLT1A"
13557..53516
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Qc3H10"
complement(13836..13904)
/rpt_family="LIMC2"
complement(13939..14597)
/rpt_family="LIMD1"
complement(14609..14916)
/rpt_family="L1"
complement(15018..15180)
/rpt_family="L1"
15186..15590
/rpt_family="MLT1B"
```





|||||  
Db 8770 CAGATAGAGACCCAAATAATCCCTTCATGCCAGTCCCTGGACCATCTGGTGGACTT 8711  
QY 360 CTGAGGCTGGCCACCCAGCCCTTGTCTGCTTAAACACACTGGGACTCGAAGTCAGA 419  
Db 8710 CTCAGGCTGGCCACCCAGCCCTTGTCTGCTTAAACACACTGGGACTCTGAAGTCAG 8651  
QY 420 GCTCCGTGTGATCAGGAAGGGTGTGTAGGAGAGGCGCTCCAGGCTCTGCCAGACA 479  
Db 8650 AGCTCGTGTGATCAGGAAGGGTGTGTAGGAGAGGCGCTCCAGGCTCTGCCAGACA 8591  
QY 480 TCATGCTCAGGATCTCAAGGAGGGCTGAGGGTCCCTTAAGACCCCACTCCCGTGACCCAA 539  
Db 8590 TCATGCTCAGGATCTCAAGGAGGGTGTAGGGTCCCTTAAGACCCCACTCCCGTGACCCAA 8531  
QY 540 CCCCCACTCCAATGCTCACTCCCGTGACCAACCCCTCTTCATTGTTCATTCACACCC 599  
Db 8530 CCCCCACTCCAATGCTCACTCCCGTGACCAACCCCTCTTCATTGTTCATTCACACCC 8471  
QY 600 ACCCCACATCCCCACCCATCCCTCAACCCCTGATGCCCATCCGCCAGGATCCGGTCCC 659  
Db 8470 ACCCCACATCCCCACCCATCCCTCAACCCCTGATGCCCATCCGCCAGGATTCACCC 8411  
QY 560 TCACCCCCACCCACCCACCCCTCCACCCCTCCAGCCCGCCAGGAGATCCGGTCCC 719  
Db 8410 TCACCCCCACCCACCCACCCCTCCACCCCTCCAGCCCGCCAGGAGATCCGGTCCC 8351  
QY 720 GCCAGGAACATCCGGTGGCCGGTGTGACGCACTGACTTGCATTTGGGGCAGAG 779  
Db 8350 GCCAGGAACATCCGGTGGCCGGTGTGACGCACTGACTTGCATTTGGGGCAGAG 8291  
QY 780 AGAAGCAGGTTTCCATTTCTGAGGAGCGGCTAGAGTTGGCCGGAAGAACCTTGACCCAG 839  
Db 8290 AGAAGCAGGTTTCCATTTCTGAGGAGCGGCTAGAGTTGGCCGGAAGAACCTTGACCCAG 8231  
QY 840 GCTCTGTGAGGAGCAAGGTGAGAGGCTGAGGAGGACTGAGACCCCGCCACTCCAAAT 899  
Db 8230 GCTCTGTGAGGAGCAAGGTGAGAGGCTGAGGAGGACTGAGACCCCGCCACTCCAAAT 8171  
QY 900 AGAGAGCCCAATATTCCAGCCCGCCCTTGTCTGCGAGCCCTGGCCACCCCGGGAAG 959  
Db 8170 AGAGAGCCCAATATTCCAGCCCGCCCTTGTCTGCGAGCCCTGGCCACCCCGGGAAG 8111  
QY 960 ACCTCTCAGCCTGGGTGCCCCCAGACCCCTGCTTCCAAAAGCCTTGAGAGACACACAGTT 1019  
Db 8110 ACCTCTCAGCCTGGGTGCCCCCAGACCCCTGCTTCCAAAAGCCTTGAGAGACACACAGTT 8051  
QY 1020 CTTCTCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAG 1079  
Db 8050 CTTCTCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAG 7991  
QY 1080 GGCAGGGCACAGGCTCTGCAGAGCATCAAGATCAGACCCCAAGAGGGAGGCTGTGGCC 1139  
Db 7990 GGCAGGGCACAGGCTCTGCAGAGCATCAAGATCAGACCCCAAGAGGGAGGCTGTGGCC 7931  
QY 1140 CCAAGAGCTCAGCTCCAATCCCACTCCCACTCCCACTTCGCAATTCCTCCACCCCAAC 1199  
Db 7930 CCAAGAGCTCAGCTCCAATCCCACTCCCACTTCGCAATTCCTCCACCCCAAC 7871  
QY 1200 CCCCCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 1259  
Db 7870 CCCCCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 7811  
QY 1260 CACCTCCAGCCCGCCAGCAGCCCAACCCCTTCTGCCACTTCAGCTCAGCTCAGCTCAGCT 1319  
Db 7810 CACCTCCAGCCCGCCAGCAGCCCAACCCCTTCTGCCACTTCAGCTCAGCTCAGCTCAGCT 7751  
QY 1320 CCCCAGCCTCATCTCTCATGTGCCCCCACTCCCATCGCTCCCGCCCAATCTTGGCAGAAATC 1379  
Db 7750 CCCCAGCCTCATCTCTCATGTGCCCCCACTCCCATCGCTCCCGCCCAATCTTGGCAGAAATC 7691  
QY 1380 CGGTTGGCCCTGCTCTCAACCCAGGGAAGCCCTGTGAGCCCGATGTGAAACCACTGAC 1439  
|||||

Db 7690 CGGTTTGGCCCTGCTCTCAACCCAGGGAAGCCCTGGATGGCCCGCATGTGAACCACTGAC 7631  
QY 1440 TTGAACCTCACAGATCTGAGAGAAGCCAGGTTTCAATTAATGTTCTGAGGGCGGCTTGA 1499  
Db 7630 TTGAACCTCACAGATCTGAGAGAAGCCAGGTTTCAATTAATGTTCTGAGGGCGGCTTGA 7571  
QY 1500 GATCCACTGAGGGGATGGTTTGTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAG 1559  
Db 7570 GATCCACTGAGGGGATGGTTTGTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAG 7511  
QY 1560 GACTGAGGAGGACACACCCAGGTAGATGGCCCAAAATGATCCAGTACACCCCTGCT 1619  
Db 7510 GACTGAGGAGGACACACCCAGGTAGATGGCCCAAAATGATCCAGTACACCCCTGCT 7451  
QY 1620 GCCAGCCCTGGACACCCCGCCAGGACAGATGCTCAGCTGAGACACCCCGCTCCCGTC 1679  
Db 7450 GCCAGCCCTGGACACCCCGCCAGGACAGATGCTCAGCTGAGACACCCCGCTCCCGTC 7391  
QY 1680 CCACTGCCACTTAACCCACAGGCAATCTGTAGTCAATAGCTTATGTAGCCGGGCAAGGT 1739  
Db 7390 CCACTGCCACTTAACCCACAGGCAATCTGTAGTCAATAGCTTATGTAGCCGGGCAAGGT 7331  
QY 1740 TGCTCAGGAGA-GGCAGGGCCAGGCAATCAAGGTCCA-GCATCCCGCCGCAATAGGTC 1797  
Db 7330 TGCTCAGGAGAAGGGCCAGGCAATCAAGGTCCAAGGTCCAAGGTCCAAGGTCCAAGGT 7271  
QY 1798 AGGACCTGGAGGGAACCTGAGGTTCCCAACCCACACACCTGCTCTCTCATCTCCACGCC 1857  
Db 7270 AGGACCTGGAGGGAACCTGAGGTTCCCAACCCACACACCTGCTCTCTCATCTCCACGCC 7211  
QY 1858 ACCCCACTCAGATTCACATACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 1917  
Db 7210 ACCCCACTCAGATTCACATACCTTACCTTACCTTACCTTACCTTACCTTACCTT 7151  
QY 1918 CTGTCACCCACGAGAGCCAGGGAATGGGCGCAGGCACTCGGATCTTACGCTCCCAT 1977  
Db 7150 CTGTCACCCACGAGAGCCAGGGAATGGGCGCAGGCACTCGGATCTTACGCTCCCAT 7091  
QY 1978 CAGGCTCTGATGAGGGAAGGGCTTGAACAGGGCTCAGGGGAGCAGAGGAGGGGCC 2037  
Db 7090 CAGGG-CTGATGAGGGAAGGGCTTGAACAGGGCTCAGGGGAGCAGAGGAGGGGCC 7032  
QY 2038 TACTGGAGATGAGGAGGCTCAGAGGAGCCAGACACCTTAGACACCCAGCCCTCTGTCT 2097  
Db 7031 TACTGGAGATGAGGAGGCTCAGAGGAGCCAGACACCTTAGACACCCAGCCCTCTGTCT 6972  
QY 2098 GAGACTGAGGCTGCCACTTCTGSCCTCAAGAATCAGAACGATGGGACACTCAGAT-GCAT 2156  
Db 6971 GAGACTGAGGCTGCCACTTCTGSCCTCAAGAATCAGAACGATGGGACACTCAGATAGCAT 6912  
QY 2157 GGGGTGGACCCAGGCTGCAAGGCTTACGCGGAGGAGGAGGAGGAGGACTCAGGGGA 2216  
Db 6911 AGGGGTGGACCCAGGCTGCAAGGCTTACGAGAGGAGGAGGAGGAGGAGGACTCAGGGGA 6852  
QY 2217 CTTGGAATCCAGATCAGTGTGGACCTCGGCTCAGAGAGTCCAGGGCAGCGTGGCCACA 2276  
Db 6851 CTTGGAATCCAGATCAGTGTGGACCTCGGCTCAGAGAGTCCAGGGCAGCGTGGCCACA 6792  
QY 2277 TATGGCCCATATTTCTGCTGCTTTCAGGTGACAGGACAGCTGCTGTGAGAGTGG 2336  
Db 6791 TATGGCCCATATTTCTGCTGCTTTCAGGTGACAGGACAGCTGCTGTGAGAGTGG 6732  
QY 2337 GGCTCAGGTTCAACAGAGGAGGAGTTCAGAGTCCATATGGCCCAAGATGTGCCCTT 2396  
Db 6731 GGCTCAGGTTCAACAGAGGAGGAGTTCAGAGTCCATATGGCCCAAGATGTGCCCTT 6672  
QY 2397 CATGAGGACTGGGATATCCCGGCTCAGAAAGAGGACTCCACACAGTCTGGCTGTCC 2456  
Db 6671 CATGAGGACTGGGATATCCCGGCTCAGAAAGAGGACTCCACACAGTCTGGCTGTCC 6612  
QY 2457 CTTTGTAGTACTTATAGGGGACCATCAGGATGGCGGTATGTTCCATTTCTCATTGT 2516  
Db 6611 CTTTGTAGTACTTATAGGGGACCATCAGGATGGCGGTATGTTCCATTTCTCATTGT 6552  
|||||





```
|||||
Db 4395 GGGGTCCAAAGGGCCCTCGCTGAAACACAGCTATGTGAAAGCCCTTGTAGTATGTGATCAAGG 4336
QY TCAGTGAAGAGTTCGGTCTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGG 4795
Db TCAGTGAAGAGTTCGGTCTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGG 4276
QY AAGAGGAGTCTGAGCATGAGTTGACGCCAAGCCAGTGGAGGGGAGCTGGGCCAGTGC 4855
Db AAGAGGAGTCTGAGCATGAGTTGACGCCAAGCCAGTGGAGGGGAGCTGGGCCAGTGC 4216
QY ACCTTCAGGGCCGCTCCAGCAGCTTCCCTGCTCGTGACATGAGGCCCATCTTC 4915
Db ACCTTCAGGGCCGCTCCAGCAGCTTCCCTGCTCGTGACATGAGGCCCATCTTC 4156
QY ACTCTGAAGAGAGCGGTGCTGCTAGTAGTGGTTCTGTTCTATTTGGTGACTTGG 4975
Db ACTCTGAAGAGAGCGGTGCTGCTAGTAGTGGTTCTGTTCTATTTGGTGACTTGG 4096
QY AGATTTATCTTTGCTCTTTTGAATTTGTTCAAAATGTTTTTTTAAAGGATGGTTGAA 5035
Db AGATTTATCTTTGCTCTTTTGAATTTGTTCAAAATGTTTTTTTAAAGGATGGTTGAA 4036
QY TGAACITCAGCATCCCAAGTTTATGAATGACAGCAGTACACAGTCTGTGTATAGTTT 5095
Db TGAACITCAGCATCCCAAGTTTATGAATGACAGCAGTACACAGTCTGTGTATAGTTT 3976
QY AAGGTAAGAGTCTGTGTTTTTATTCAGATTGGGAATTCATCTATTTTGTGAATGGG 5155
Db AAGGTAAGAGTCTGTGTTTTTATTCAGATTGGGAATTCATCTATTTTGTGAATGGG 3916
QY ATATAACACAGCTGGAATAAGTACTTTAGAAATGTGAAAATGACAGTAAATAGATGA 5215
Db ATATAACACAGCTGGAATAAGTACTTTAGAAATGTGAAAATGACAGTAAATAGATGA 3856
QY GATAAAGAACTAAAGAAATTAAGAGATAGTCAATCTTGCTTATACCTCAGCTATTCT 5275
Db GATAAAGAACTAAAGAAATTAAGAGATAGTCAATCTTGCTTATACCTCAGCTATTCT 3796
QY GTAAATTTTAAAGATATATGATACCTGATTTCTTGCTTCTTTGAGAAATGTAAGA 5335
Db GTAAATTTTAAAGATATATGATACCTGATTTCTTGCTTCTTTGAGAAATGTAAGA 3736
QY GAAATTAATCTGAATAAAGAAATTTCTTCTGCTCAGTGGCTTTTCTTCCATGCACT 5395
Db GAAATTAATCTGAATAAAGAAATTTCTTCTGCTCAGTGGCTTTTCTTCCATGCACT 3676
QY GAGCATCTGCTTTTGAAGGCCCTGGGTTAGTAGTGAGATGCTAAGGTAAGCCAGCT 5455
Db GAGCATCTGCTTTTGAAGGCCCTGGGTTAGTAGTGAGATGCTAAGGTAAGCCAGCT 3616
QY CATACCCACCATAGGCTGCTAGAGTCTAGAGCTGAGTCACTGATCCAGTGGTGGCAAG 5515
Db CATACCCACCATAGGCTGCTAGAGTCTAGAGCTGAGTCACTGATCCAGTGGTGGCAAG 3556
QY ATGTCTCTTAAGATGATAGGAAAAGTACAGAGAGGCTGAGGCTGCGGGCTCGGGTGA 5575
Db ATGTCTCTTAAGATGATAGGAAAAGTACAGAGAGGCTGAGGCTGCGGGCTCGGGTGA 3496
QY GAGTGGTGGAGTGTCAATGCCCTGAGTGGGGCAATTTGGGCTTTGGGAACTGCAGTTC 5635
Db GAGTGGTGGAGTGTCAATGCCCTGAGTGGGGCAATTTGGGCTTTGGGAACTGCAGTTC 3436
QY CTTCGGGGGAGCTGATGTTGAATGATCTTGGTGGATCC 5674
Db CTTCGGGGGAGCTGATGTTGAATGATCTTGGTGGATCC 3397

RESULT 4
AF134576 AF134576 49375 bp DNA PRI 06-APR-1999
LOCUS Homo sapiens MAGE10 and undefined MAGE genes, complete sequence.
DEFINITION AF134576
ACCESSION AF134576
```

94567140  
AF134576.1 GI:4567140  
HTG.  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
Sequence of three cosmids of human Xq28 (16c3, 3g5, 3h5) containing  
MAGE genes  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 49375)  
AUTHORS  
Gloeckner, G., Rosenthal, A. and Schattevoy, R.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (12-MAR-1999) Genome Analysis, Institute for Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
FEATURES  
Location/Qualifiers  
source  
1..49375  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="Xq28"  
/chromosome="X"  
/clone="cosmids 16c3; 3g5; 3h5"  
misc\_feature  
1..10000  
/note="incomplete at 3' end"  
repeat\_region  
/note="Region: Cosmid 16c3"  
complement(10..202)  
/rpt\_family="MER93"  
repeat\_region  
/evidence-not\_experimental  
289..372  
/rpt\_family="LIM1"  
repeat\_region  
/evidence-not\_experimental  
446..726  
/rpt\_family="LIM1"  
repeat\_region  
/evidence-not\_experimental  
528..43599  
/note="Region: Cosmid 3g5"  
repeat\_region  
731..4392  
/rpt\_family="LIMC3"  
repeat\_region  
/evidence-not\_experimental  
complement(4393..4696)  
/rpt\_family="MER2"  
repeat\_region  
/evidence-not\_experimental  
complement(4697..4755)  
/rpt\_family="LIMC2"  
repeat\_region  
/evidence-not\_experimental  
complement(4756..4803)  
/rpt\_family="MER2"  
repeat\_region  
/evidence-not\_experimental  
4805..4847  
/rpt\_family="LIMC3"  
repeat\_region  
/evidence-not\_experimental  
4837..5414  
/rpt\_family="LIMC5"  
repeat\_region  
/evidence-not\_experimental  
5566..6545  
/rpt\_family="LIM4"  
repeat\_region  
/evidence-not\_experimental  
complement(6505..6963)  
/rpt\_family="MIR"  
repeat\_region  
/evidence-not\_experimental  
7671..7697  
/note="111 with 85% homology to 112"  
repeat\_region  
/rpt\_type-inverted  
/evidence-not\_experimental  
7814..7981  
/note="homology = 66.10%, score = 48, counts = 28"  
repeat\_region  
/rpt\_type-tandem  
/rpt\_unit-ccccac  
misc\_feature  
7814..8708  
/evidence-not\_experimental  
/note="GC score = 44.00 (895bp)"

```
/note="Region: GC content"
/region="not_experimental"
misc_feature
8294..8511
/note="CpG island score = 0.74, GC = 66.50%, CpGs = 17"
/note="Region: CpG island"
/region="not_experimental"
repeat_region
8854..8880
/note="I12 with 85% homology to I11"
/rpt_type="inverted"
/region="not_experimental"
gene
9050..11121
/region="MAGE10"
/note="I1..2072 of gb|U10685|HSU10685 Human
MAGE-10 antigen (MAGE10) gene, complete cds. 6/95 P = 0.0
S = 9975"
9936..10064
/region="MAGE10"
/note="GC score = 5.60 (129bp)"
/region="not_experimental"
misc_feature
10412..10533
/region="MAGE10"
/note="GC score = 8.80 (122bp)"
/region="not_experimental"
misc_feature
11591..49375
/note="Region: Cosmid 3h5"
/region="not_experimental"
repeat_region
12154..12482
/rpt_family="L1M4"
/region="not_experimental"
repeat_region
12570..12980
/rpt_family="Alus"
/region="not_experimental"
repeat_region
14281..14525
/rpt_family="MIR"
/region="not_experimental"
repeat_region
16047..17011
/rpt_family="L1R5"
/region="not_experimental"
repeat_region
17625..17747
/rpt_family="L1ME1"
/region="not_experimental"
repeat_region
17800..17815
/rpt_family="L1M4"
/region="not_experimental"
repeat_region
17843..18317
/rpt_family="L1M4"
/region="not_experimental"
repeat_region
20010..20829
/rpt_family="L1MC/D"
/region="not_experimental"
repeat_region
20859..21195
/rpt_family="MLT2E"
/region="not_experimental"
repeat_region
21198..21305
/rpt_family="L1MC/D"
/region="not_experimental"
repeat_region
21306..21496
/rpt_family="MLT1B"
/region="not_experimental"
repeat_region
21844..21992
/rpt_family="L1MC/D"
/region="not_experimental"
repeat_region
22019..22814
/rpt_family="L1MC/D"
/region="not_experimental"
repeat_region
22796..23134
/rpt_family="L1M4"
/region="not_experimental"
repeat_region
23127..23300
```

```
/rpt_family="L1MD3"
/region="not_experimental"
repeat_region
23271..23837
/rpt_family="L1MC5"
/region="not_experimental"
repeat_region
24267..25020
/rpt_family="L1M4"
/region="not_experimental"
misc_feature
26071..28617
/note="GC score = 31.80 (2547bp)"
/region="Region: GC content"
/region="not_experimental"
repeat_region
26655..26726
/note="homology = 79.20%, score = 36, counts = 12"
/rpt_type="tandem"
/rpt_unit="catccc"
/region="not_experimental"
repeat_region
27233..27358
/note="homology = 65.90%, score = 34, counts = 21"
/rpt_type="tandem"
/rpt_unit="ccccca"
/region="not_experimental"
misc_feature
27403..27510
/note="CpG island score = 1.29, GC = 61.10%, CpGs = 9"
/region="Region: CpG island"
/region="not_experimental"
repeat_region
27829..27906
/note="homology = 79.50%, score = 40, counts = 13"
/rpt_type="tandem"
/rpt_unit="tccccca"
/region="not_experimental"
repeat_region
28990..29073
/note="homology = 69.00%, score = 26, counts = 14"
/rpt_type="tandem"
/rpt_unit="ccccca"
/region="not_experimental"
misc_feature
30461..30620
/note="GC score = 5.00 (160bp)"
/region="Region: GC content"
/region="not_experimental"
repeat_region
30896..30958
/note="homology = 82.50%, score = 20, counts = 3"
/rpt_type="tandem"
/rpt_unit="aggcttagagatcacccagg"
/region="not_experimental"
gene
31110..31778
/region="undefined MAGE"
/note="I1..185 of gbp|D32077|HUMMAGEC_1 Homo-sapiens Human
mRNA for MAGE-41 protein, complete P = 6.0e-95 S =
592--193..283 of gbp|D32076|HUMMAGEB_1 Homo sapiens Human
mRNA for MAGE-6 protein, complete cds; P = 1.5e-80 S = 356"
Query Match 48.9%; Score 2777.2; DB 42; Length 49375;
Best Local Similarity 74.9%; Pred. No. 0;
Matches 4355; Conservative 0; Mismatches 1109; Indels 353; Gaps 56;
QY 1 CCGGGGACCACTGGCATCCCTCCCTACACCCCAATCCCTCTTACGCCACCC 60
DB 27297 CCATGAGAACCCCATCTCTCCAGGCCCATCTATCTATCCACCCACCCACCCCT 27356
QY 61 ATCCAAACATCTTCACGCTCACCCAGCCCAAGCCAGCAATCCGGTTCCACCCCTG 120
DB 27357 CATCTACCTACCTACCCCTATCCCTATCTGTAGTAATCCGGTTTCGCCCTG 27416
QY 121 CTCTCAACCCAGGGAAG-CCAGGTGCCAGATGTGAGCCACTGACTTGACATTAGTG 179
DB 27417 CTTTCAACCGAGGAATCCCGGGTGGCCGGATGTGACACCACTGACTTG-CCGTGGGG 27475
QY 180 GTTAGAGAGACGAGGTTTCGGTCTGAGGGGGGGCTTGTAGATCGGTGGAGGAGCGG 239
DB 27476 GTCAAGAGAGACGAGTTCTCGTCTGAGTGGCGGCTTGAGATTGGTGGAGGAGTGG 27535
QY 240 GCCCA-GCTCTGTAAAGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACC 298
```

Db 27536 GTCCAGGCTCTGTAGGAGGCAAGGTGAGCTCTAAGGAGGAGCTCAGAGGCCCCAC 27595  
Qy 299 CCAGATAGAGGACCCCAATAATCC-----CTTCATGCGAGTCTCTGGACCATCTGG 349  
Db 27596 CCAGATAGAGGCCCCCAATAATCCAGCACTACTCTGCTGCGAGCCCTAACCACCTGG 27655  
Qy 350 TGGTGGACTTCTAGGCTGGGCGACCCCGACCCCT-----TGTCTTAAACCA 400  
Db 27656 GGGCGGACTTCTAGCTTGGGCGACCCCTGCAACCCCTGCGCTTGAAGCG 27715  
Qy 401 CTGGGGACTC--GAATCAGAGCTCGTGTGATCAGGAGGAGGCTGCTAGGAGAGGGCAG 459  
Db 27716 CAAGGAGCTCTGGAGCTAGAGCTTCTGTGACCAAGGCGAGGCTGGTTAGGAGAGGGCAG 27775  
Qy 460 CGTCCAGGCTCTGCGAGACATCATGCTCAGGATTTCTCAAGGAGGCTGAGGCTCCCTAAG 519  
Db 27776 TGGCCAGGCTCTGCCAGGATCAATCAGGACCTTAAGTGAAGGCTGAGGCGCCCCAC 27835  
Qy 520 ACCC---CACTCCCGTGACCCCAACCCCACTCCAATGCTCACTCCCGTGACCCCAACCC 576  
Db 27836 CCCCATTCCTACCCCAACCCCACTCCCATCCCATCCCATCCCATCCCATCCCA 27895  
Qy 577 TCTTCATGTGATTCGAACCCCAACCCCACTCCCATCCCATCCCATCCCATCCCAT 636  
Db 27896 CCAGAACCCCTATCCCCCAACCAAGAACCTATCTGCGCACTCCCAACCACTCCATCC 27955  
Qy 637 CCATCCGCCCA--GCCATTCACCCCTACCCCAACCCCAACCCCAACCCCACTCCCA 695  
Db 27956 TACTCCCACTTCCATCTCCCAACCCCTGACCTCTCTCTCACTCCCAACCCCA 28015  
Qy 696 CCACCCAGGAGATCCCG-----TTCCCGCGAGGAACATCCCGGTGCCG 742  
Db 28016 ACACCTGGCAGATTCGGTTCTGCTCTGCTTTTAAACCCAGGGAATCCCTGGGTGACCA 28075  
Qy 743 GATGTGAGCCCACTGACTTGGCATTTGGGCGAGAGAGAGAGGAGTTTCATTTCTGAG 802  
Db 28076 GATGTGGTGCCTGCTGTGCATTTGAGGTGCGAGAGAGAGGCGCTCGCTCTCAG 28135  
Qy 803 GGACGGCTAGAGTTTCGGCGGAGGAACCTGACCCAGGCTCTGTGAGGAGGCAAGGTGAG 862  
Db 28136 GGGCAGC--TGGAGATCAGCTGAGGCGAGCTGGCCCTGGCTCTGTGAGGATGCAAGGTGAG 28194  
Qy 863 AGGCTGAGGAGGAGTGAAGACCCCGCCCACTCCAATAGAGAGCCCAATATTCAG-- 921  
Db 28195 ATGCTGAGGAGGAGTGAAGGATATCCACCCCTGCTGAGTGGACCCCAATATCCAGTG 28254  
Qy 921 -CCCGCCCTTCTGCTGCGAGCCCTGGCCCAACCCCGGGAAGAGCTCTACGCTGGGCTGCC 979  
Db 28255 CCACCTCTCTGCTGCTAGCTCTGACCACTCCAGGCGAGGACTCTTAGGCTGGGCCAC 28314  
Qy 980 CCAGAGCCCTCTCCAAAGCCTTGAGAGACACAGGTTCTTCTCCCAAGCTCTGGAA 1039  
Db 28315 CCCAGTCCCGCACCGTTAAGCCGAGGGGA-----CTCAGGAG 28353  
Qy 1040 TCAGAGTTGCTGTGACACGGCAGGAGTGGTTAGGAGAGGCGAGGCGACAGGCTCTGCC 1099  
Db 28354 ACAGAGCTTGGTATGACACGGCAGGAGTGGTTAGGAGAGGACAGCTCCCGAGCTCTGCC 28413  
Qy 1100 AGGCATCAAGATCAGACACCAAGAGGAGGCTGTGGGCGCCCAAGAGCTGCACTCCAATC 1159  
Db 28414 AGGAACACGCTCAGGAACCTAAGGGAAGCTGAGGCTACCC-----28457  
Qy 1160 CCCACTCCACCCCATTCGCATTCCTATTCCTCCCAACCCCACTCTCTCAGCTACAC 1219  
Db 28457 -----CAC 28459  
Qy 1220 CTCCACCCCACTCTCTACTCTGCTGCTGACCTGACCAACCACTCCAGCCCAAGCAGCC 1279  
Db 28460 CCCAATCTATTTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 28514  
Qy 1280 AGCCCAACCCCTTCTGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339

Db 28515 CACCCCATCTCTACCGGCACTCTATCCCATCATCCCCA-----28554  
Qy 1340 TGTGCCCACTCCCATCGCTCCCGCATTTGGCAGAAATCCGGTT--TGCCCCCTGCTCTCA 1398  
Db 28554 -----CCCCATCTCTGGCAGAAATCCGATTCGCCCTGATTCA 28592  
Qy 1399 ACCCAGGAAGCCCTGGTAGGCCCGATGTAAACCACTTGACTTGAACCTTCACAGATCTGA 1458  
Db 28593 ACCCAGGAAGCCCTGGAGGCCCGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 28652  
Qy 1459 GAGAGCCAGGTTTAAATGTTCTGAGGGCGGCTTGAGATCCACTGAGGGGAGTGG 1518  
Db 28653 GAGATCAAGG-----GCATGGTTCTGAGAAGCGACTGAGATCAGCAGAGGGGAATGG 28706  
Qy 1519 TTTTATGCTGTGAGGAGCAAGGTGAGA--TGTCTGAGGGAGGACTGAGAGGGCACACAC 1577  
Db 28707 GCCCGGGCTCTGTGAGGAGCAAGGTGAGACCCCGAGGAAGATGAGGAAGCCCTCAC 28766  
Qy 1578 CCCAGGTAGATGGCCCCCAAAATGATCCAGTACCAACCCCTGTGCGAGCCCTGGACACCC 1637  
Db 28767 CCA--GATAGAGAACCCCAAAATATCCAGTACTACTTGTGCGAGCCCTGGACAC-- 28823  
Qy 1638 GGCCAGGACAGATGCTCAGCTGGACACCCCGCTCCGCTCCCTGCCACTGCCACTTAACCA 1697  
Db 28823 --CCAGGCGAGACTTCTCAGGCTGGACCTTCCC-----CCCTCCCACTGCCACTTAAGCA 28877  
Qy 1698 CAGGCAATCTGTAGTCTATAGCTT--ATGTGACCGGGGAGGGTTGGTCAAGGAGAGCAGG 1756  
Db 28878 CNAAGGACTCTGGAGTCAGACCTTGGTGTGACCGGGAAGGCGCGTCAAGGAGAG-- 28934  
Qy 1757 GCCCAGGATCAAGTCCAGCATCCCGCGCATTAGGGTCAAGGACCCCTGGAGGGAAT 1816  
Db 28934 -----GCAGGGCGAGGCTCTGTCAAGCATCAAAATCAGGACCCCTGAGAGAAAT 28984  
Qy 1817 GAGGTTTCCCCACCCACACCTGCTCTCTCATCTCCACCGCCACCCCACTCACTTCCCAT 1876  
Db 28985 GAGGCGCCCGCCCAACCCCACTTACCCATCCCTAACCCCATACCCCACTCTACTTGCAT 29044  
Qy 1877 ACCTTACCTTACCCCA-----ACCTCATCTTGTGAGAA-----TCCTGCTGTC 1922  
Db 29045 CCCAGCCCATCCCAACACCTTACCCCATCTTGGCAGAAATCTGTTCTTCCCTGCGATC 29104  
Qy 1923 AACCCAGGAAGCCACGGAATGCGGCGAGCACTCGGATCTTGACGCTCCCATCCAGG 1982  
Db 29105 AACCCAGGAAGCCCAAGGATGACAGACAGCACACCACTTCTGACGCTCCACATCCAGG 29164  
Qy 1983 GTCTCATGAGGAGGAGGCT-----TGAACAGGCGCTCAGGGGAGCAGAGGAGG-- 2034  
Db 29165 GCTGAAGGAGGGAAGGCTTAGTATCATGAGCAGGCGCTCAGGGAGTCTCTGCTCCTC 29224  
Qy 2034 --GCCCTACTGCGAGATGAGGAGGCGCTCAGAGGACCCAGCACCTTAGA-----C 2082  
Db 29225 AAGCCCTGCTGGAGTAAGGAGGCGCTCAGGGAACCCAGGCTCTCAGGATAGGGGTCC 29284  
Qy 2083 ACCGACCCCTCTCTGAGACTGAG--GCTGCCACTTCTGGCCTCAAGATCAGAACGATG 2140  
Db 29285 ACTCAACCCCTCTCTGAGACTGAGGCGCTCTCTTTCATCTCTCGGAATCAGAGGATG 29344  
Qy 2141 GGGACTCAGATTCATGGGGTGGGACCCCGCTTGAAGGCTTACGGGAGGAGAGGA 2200  
Db 29345 GAGACTCAGTCAGCAGGAGGTGGGGCCCAACCTCTGCCAGATCAAGGAGAGGAAGA 29404  
Qy 2201 GGGAGGACTCAGGAGGCTTGAATCCAGATCAGTGTGAGACCTCGCCCTGAGAGGTCCA 2260  
Db 29405 GGGAGGACTCAGGAGTACCTTTGAGTCCAGAACATGGGAGCCTTTGCCCTGGAGGTCCA 29464  
Qy 2261 GGGCAGGTGGCCACATATGGCCCATATTTCTGCTCTTTGAGGTGAC-----AGCAG 2316  
Db 29465 GTGACAGTGGCCACCTGTAGCCCATGTTGCTGACCTTCTGGGTGACAAAGAGGAG 29524  
Qy 2317 AGCTGTGTCTGAGAGTGGGGCTCAGTCAACAGAGGAGGAGTTCAGGATCCATAT 2376  
Db 29525 GGCTGTGTCTGAGCAGTGGTGTGACTCAGTCAAGGAGGAGGAGTCCAGCATCTGAG 29584

QY 2377 GCGCCAGATGTGCCCTTCTCAGGAGTGGGATATCCCCGCTCAGAAAGAGGAC 2436  
DB 29585 GCGCCAAATGTGCCCCATTCATGAAGTTGGGATA-CCTTGGCTCAGAAAGAGGAC 29643  
QY 2437 TCCACAGAGTGTGCTGCCCTTTTAGTAGCTAGGGGACCAGATCAGGATGCGG 2496  
DB 29644 CCCACAGAGTGTGCTGCCCTTTTAGTAGCTAGGGGACCAGATCAGGATGCGG 29703  
QY 2497 TATGTTCCATTTTCACTGTACACAGGAGAAATTTGGGGGCCCCCAGGAGATGGG 2556  
DB 29704 TATGTGCCAACCTCATTTGTGCCACAGAAAGATTTAGAGAGCCCTCAGGATGGG 29763  
QY 2557 TCTTGGGTAAAGGGGATGTCTACTCATGTCTCAGGAAATTTGGGGTTAGGAAGACAG 2616  
DB 29764 TCTTGCAGTAAAGGGAGCTATCTGCTCATCTCAGGGGTTTTCAGGTTAGGAATGCGAG 29823  
QY 2617 GCGTGGCAGGAATAAGATGAGTGAGACAGACAAAGGCTATTTGAATCCACACCCAGAA 2676  
DB 29824 GCGCCATCAGC-ATGAAGAGTAAACCCACAGGAGGCCATAGAAACACTCACCCAGAA 29879  
QY 2677 CCAAGGGGTGAGCCCTGGACACCTCA-CCAGGATGTGGCTTCTTTTTCAC 2727  
DB 29880 CCAAGGGGTATACCTTGGACACCCCATTTGGGGGTACAGAGATGTAGC-TCCATCTCAT 29938  
QY 2728 TCTGTTTCCAGATCTGGGGCAGGTGAGGACCTCATTTCTCAGAGGTTGACTCAGGTCAAC 2787  
DB 29939 TCTGTTTTCAGATCTCGGGAGGTGAGGAATTTGTTCTCCGAGGATGACTCAGGTCAAC 29998  
QY 2788 GTAGGGACCCCATCTGGTCTAAGACAGAGCGGTCCAGGATCTGCCATGCTTTCGGT 2847  
DB 29999 ACAGGGCCCCCATCTGGTGTATAGACAGAGTGGTCCAGGATCTGTCACTAGTTCCGGT 30058  
QY 2848 GAGGAACATGAGGGAGGACTGAGGGTACCCAGGACAGAACACTGAGGAGACTGCACA 2907  
DB 30059 GAGGAACATGAGGGACGATGAGGGACCTTTGGGCCAGAACAGATGAGGACCTCAG 30118  
QY 2908 GAAATAGCCCTGCCCTGTGTACCCACAGAGAGCATGGGCTGGGCGCTGCCCCAGGT 2967  
DB 30119 GAAATAGCCCTGCCCTGTGTACCCACAGAGAGCATGGGAGGCTGTCTGCTGAGT 30178  
QY 2968 CTTTCC-GTTATCCTGGGATCATTGATGATCAGGAGCGGGAGGCTTGTCTGAGAGG 3025  
DB 30179 CCCCCCCTTACCTGGGATCTGTTGTGTCAGGAGTGGGAGGCTTTGTC-GAGGGGT 30237  
QY 3026 CTGCGCTCAGGTGAGTGGAGCGTCCAGGCGCTCCAGGAGTCAAGGTGAGGACCA 3085  
DB 30238 CTGCACTCAGGTGAGTGGAGCGTCTTAGGCGCTGCCAGGACAGAGTAAAGACGA 30297  
QY 3086 AGCGGGCACCTCACCCAGGACATTAATTTCCAAATGAAATTTGATATCTCTTGTGCCCT 3145  
DB 30298 AGCAGGTTCTCACCCAGGACATGAATTTCCAAATGCAATTTTCACTCTCTTCTGTCT 30357  
QY 3146 TCCCCA-AGGACCTAGGACGTTGGCCAGATGTTTGTCCCTCTCTTCTTCTTCTCT 3204  
DB 30358 TCCCAAGGAGGCTGGGACGTTGGCCAGATGAGTCTCTCTCTCTCTCTCTCTCTCT 30414  
QY 3205 TATCATGAGTGTGAACCTTTGATTTGATTTCTCAGACCAAGAAAGGCGAGGATCCAGG 3264  
DB 30415 TATCAGGATGTGAGCTTTAATCTGAGTTCTCAGGCCAGAAAGGTTGGATCCAGG 30474  
QY 3265 CCCTGCCAGGAAATATAGGGCCCTGCGTGAGACAGAGGGGGTATCCACTGTCATGA 3324  
DB 30475 CTTTGCAGGAGAAAGTGGGGCCCTGTGTGAGCAGAGGGGACCATTCACCCCAAGA 30534  
QY 3325 GAGTGGGATGTACAGAGTCCAGGCCACCTCTCTGTGTAGACTGTAGAAACCCAGGCTGT 3384  
DB 30535 GGTGGGAGACCTCACAGATTTCCAGCTTACCCTCTCTGTGTAGACTGGGGGCTTGTGCTGT 30594  
QY 3385 GCTTGGCGTGTGACCCCTGAGGGCCCTGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3444  
DB 30595 GCTTGGCAGTGTGACCCCTGAGGGCCCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3464

QY 3445 CAGTGGGCGCTTGGTCTGAGACAGATATCTCTCAGGTACAGAGCAGAGGATGTCACAG- 3501  
DB 30655 CAGTGGGCGCTTGGTCTGAGGCGCTGCCCTCAGTTCAGTACAGAGCAGAGGAGATGACAGCGT 30714  
QY 3501 GGTGTGCCAGCAGTGTATGTTTGGCCCTGAATGACACCAAGAGGCGCCACCTCCACAGGA 3560  
DB 30715 CTAGTGGCAGCAGTGAACGTTTGGCTTGAATGACACTAATAGGCCCCCATCGCCCCAGAA 30774  
QY 3561 CACATAGGACTCCACAGAGTCTGGCTCA-CCTCCCTACTGTCTAGTCTCTAGAAATGCAC 3619  
DB 30775 CATATGGGACTCCAGAGCAGTCTGCCCTCAGTCTCTCTACTGTCTAGTCTCTCAGAAATCAGC 30834  
QY 3620 CTCTGTGGCGGCTGTACCTTGAGTACCTCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3679  
DB 30835 CTCTGTGTCTGTGTATACCTGTAGGTCCTCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 30894  
QY 3680 CAGGCCAACCCAGAGC-AGGATTCCTCTGGAGGCCACAGAGGAGCAGC 3726  
DB 30895 CAGGCTGACCAAGATCACCAGGAAGCTCCAGAGGATCCCCAGGAGGCCCTTAGAGGAGCAGC 30954  
QY 3727 C-AAGGAGAAGATCTGTAACTAGGCGCTTTGTTAGAGTCTCCAAAGGTTTCACTCTCTCAGCTG 3785  
DB 30955 CAAGGAGAAGATCTGTAACTAGGCGCTTTGTTAGAGCTCCAAAGGTTTCACTCTCTCTCTCTCT 31014  
QY 3786 AGGCTCTCACACACTCCCT 3845  
DB 31015 AGGCTCTCACACT 31074  
QY 3846 CCACACT 3905  
DB 31075 CCACACT 31134  
QY 3906 ACTGCAAGCCTGAGGAGCGCTTTGAGGCCCAACAGAGGCCCTGGGCGCTGGTGTGTGTC 3965  
DB 31135 ACTGCAAGCCTGAGGAGCGCTTTGACACCCCAAGAGAGGCCCTGGGCGCTGGTGTGTGTC 31194  
QY 3966 AGGCTGCCAC-TCCT 4004  
DB 31195 AGGCTGCCACTACTGAGGAGCAGGAGGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 31254  
QY 4005 CCCTGGAGGAGTGGCCACTGCTGGGTCAACAGATCTCTCCACAGAGTCTCTCAGGAGCGCT 4064  
DB 31255 CCCTGGGAGGAGTGGCTGCTGGGTCAACAGATCTCTCTCAAGAGTCTCTCAGGAGCGCT 31314  
QY 4065 CCGCTTTTCCACTACCATCAACTTCACTCGACAGAGGACCCAGTGAAGGTTCCAGCA 4124  
DB 31315 CCGCATCCCACTGCCATGATTTCACTCTATGAGGAGCAATCCATTAAGGGCTCCAGCA 31374  
QY 4125 GCGTGAAGAGGAGGCGCCCAAGCACCTCTTGTATCTCTGGAGTCTCTTCTTCTCCGAGCAGTAA 4184  
DB 31375 ACCAAGAAGAGGAGGCGCCCAAGCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 31434  
QY 4185 TCACTAAGAGGTTGGCTGATTTGGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4244  
DB 31435 TCACTAAGAGGTTGGCTGATTTGATTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 31494  
QY 4245 CAGTCAAAAGGAGGAGTCTGAGAGTGTCTCAATAAATTAACAAAGTCTGTTTCTCTCT 4304  
DB 31495 CGGTCAAAAGGAGGAGTGTCTGAGAGGCTCATCAAAATTAACAAAGTCTGTTTCTCTCTCT 31554  
QY 4305 AGATCTTCGGCAAGGCTCTGAGTCT 4364  
DB 31555 AGATCTTCGGCAAGGCTCTGAGTCT 31614  
QY 4365 CAGACCCACCGGCT 4424  
DB 31615 CAGACCCACCGGCT 31672  
QY 4425 TGTCTGGTGTATATCAGATCATGCCCAAGAGGCTTCTCTGATAAATTTGCTCTCTCTCTCTCTCT 4484  
DB 31673 TGTCTGGT--TATCAGATCATGCCCAAGAGGCTTCTCTGATAAATTTGCTCTCTCTCTCTCTCT 31730  
QY 4485 TTCAATGAGGCGGCCCATGCT 4544



ESLFRALSNKVDLAHFLRKYRAKELVTKAEMLEIRVKNYKRCPPVIFGKASESLK  
MIFGIDVKEYPASNTYTLVTCLGSLYDGLGNQIIPFKTGLLIIVIGITAMGDSAS  
EEIWEELGVYGVYDREHTVYGEPRKLLTQDWQENYLEYRQVPGNSPNARYEFLWGP  
RALAETSIVKVLHVVRNVRVIAVPSLREALLLEEVEG"  
polya\_signal 11072..11077  
polya\_site 11091  
BASE COUNT 2330 a 3788 c 2924 g 2253 t  
ORIGIN

Query Match		46.88;	Score 2655;	DB 9;	Length 11495;
Best Local Similarity		73.68;	Pred. No. 0;		
Matches 4231;		Conservative	0;	Mismatches 1100;	Indels 416; Gaps 50;
Qy	21	CCTCCCTTACACCCCAATCCCTCCCTTTACGCGCCACCCATCCAAACATCTTCACGCTC	80		
Db	5969	CCAGCACCCCTATCTCCCAACCCCACTACCTTATGCTTCATCCCAACCCCAAC	6028		
Qy	81	ACCCCAAGCCCAAGCAGGAGAGATCCGGTTCCACCCCTGCTCTCAACCCAGGAAGC-C	139		
Db	6029	ACCACTATCCCATCCAGGTTGAATCGCATTCCTCCCTTTCTGCTTTCAACCCAGGAAGTC	6088		
Qy	140	CAGGTGCCAGATGAGCCGACATGACTTGAGCATTAGTGTGTAGAGAGAGCGAGTTT	199		
Db	6089	CAGGTTCCTGGATGTGATGCCAGTGACTTGTGCATTGGGGGTTAGAGAGAGCTAGCTTC	6148		
Qy	200	TCGGTCTGAGGGGGGCTTGAGATCGGTGGAGGAAGCGGGCCCA-GCTCTGTAAAGAGG	258		
Db	6149	TCAGTCTGACAGGAGCTTGGATTGGCAGAGGAAGCGGTCCAGGCTCTGTAGAGTGG	6208		
Qy	259	CAAGGTGACATGCTGAGGAGGAGT---GAGGACCCATTTACCCAGATAGAGAGCCCA	315		
Db	6209	CATAGTGAAGACTGAGGAGGAAGTGGGAGGCGCTCTCCACCCAGATAGAGAGCCCA	6268		
Qy	316	AATAATCC-----CTTCATGCCAGTCTGACCATCTGTTGGTGGACTTCTCAGGC	366		
Db	6269	AATAATCGGACCCCTCTCTCCAGTCTGGGCGCACCCGTTGGGGGAGTCTTGAGTC	6328		
Qy	367	TGGGCCACCCCGCCCTTGCTGTCTAAACCACTGGGGACTCGAAGTCAAGACTCCGT	426		
Db	6329	TGGGACGCCACCAACCCACTGCGCTGAAGCCGAGGAGTATGGAGTCAGACTTGGT	6388		
Qy	427	GTGATCAGGAAGGCTGCTGTAGGAGAGGCGAGCGTCAGGCTCTGCCAGACATCATGCT	486		
Db	6389	GTGATCAGTGCAGGACTGTTGGGGT-----AGGCTCTGCCAGGACATCAAGT	6436		
Qy	487	CAGGATCTCAAGAGGCTCAGGCTCCCTAAGACCCCACTCCCGTGACCAACCCCAAC	546		
Db	6437	CAGGACCTTAGGAGGCTGAGTGTCCCCCA---CCCCCATTCCTATCCCTTACCCCTT	6493		
Qy	547	TCCAATGCTCACTCCCGTGACCAACCCCTCTTCTATTGTTCATTCCAAACCCCAACCC	606		
Db	6494	TCCCATCTGCACTCC-----TACCCCATCTGTACCCCC	6527		
Qy	607	ATCCCCCAACCCATCCCTCAACCTGATGCCCATCCGCCCCAGCCATTCCACCCCTCAC	666		
Db	6528	ATTCACCCACCTGTGCCCTTATCTCCCAACCCCAACCCCAACCCCAACCCCTCCCC	6587		
Qy	667	CACCCCAACCCCAACCCCACTCCCAACCCCAACCCCAACCCCAACCCCAACCCCAAC	725		
Db	6588	CACCCCTACCTTCACTCCCATCTAGTGAGCATCCGGTTCCACCCCTGCTTTCAATCCAG	6647		
Qy	726	AAACATCCGGGTGCCGGATGTGACGCACTGACTTGGCATTTGTGGGCGAGAGAGAC	785		
Db	6648	CAAGCCCTGGGTGCCGGATGTGATGCCACTGACTTGTGAATTTAGGGTTAGAGAGAT	6707		
Qy	786	GAGGTTTCCATTCTGAGGAGCGGCTAGAGTTGGCGGAAGAACCTGACCCAGGCTCTG	845		
Db	6708	GAGTTTCTGGTCTGAAGGGTGGC-TTGAGATCGCAGAGGAAGGTGGCCAGGCTTG	6766		
Qy	846	TGAGGAGGACGCTGAGAGGCTGAGGAGGAGTCTAGGACCCCGCCCACTCCAAATAGAG	905		
Db	6767	TGAAGAGGCAAGTGAGACTCTGAGGAGGAGATTGAGGAACCCCTTATCCCTGTATAGAG	6826		

Qy	906	CCCCAAATATTCAGCCGCCCTTGTGTCAGCCCTTGCCAGCCCTGGCCACCCGCGGAGACACTCT	965		
Db	6827	TCCAGAGCCCTGGACTACC-----CGCGAGGCTCACTTCT	6862		
Qy	966	CAGCTGGGTGCCCCAGAGCCCTGCTTCAAAAGCCCTTGAGAGACACAGGTTCTTCTC	1025		
Db	6863	CAGACTGGGTGCTCCGCCACTCCGCCCC-----TTGCAACGCGTTTGTAAAGCAC	6917		
Qy	1026	CCCAAGCTCTGGAAATCAGAGTTCTGTACCCAGGCGAGGACTGTTAGGAGGCGGAGG	1085		
Db	6918	AGGGACTCTGGAGTCAGAGTTGGTGTGATCAGGAAGGCTGTTAGGAGA-GGATG	6976		
Qy	1086	GCAGAGCTCTGCCAGGATCAAGATCAGACCCCAAGAGGAGGCTGTGGGCCCAAG	1145		
Db	6977	GCCAGGCTGCCAGGAATCAAGTCAGAACCC-TGAGAGGAACCTGAGGTCGCCCAAG	7035		
Qy	1146	ACTGCACTCAATCCCACTCCACCCCATTCGCAATTCCTCCCAACCCCAACCCCAT	1205		
Db	7036	ATCCTAGTCTAAACCCCACTCCCAAC-----	7063		
Qy	1206	CTCCTCAGCTACACTCCACCCCATCCCTACTCTCTACTCCGTACCTGACCCACCT	1265		
Db	7063	-----ATCCGCTGCCATTTGCTGCTCCATTTCCCATTTCTTTCGCT	7104		
Qy	1266	CCAGCCCAAGCAGCCAGCCCAACCCCTTCTGCCACTCACCTCACTGCCCAACCCCAAC	1325		
Db	7105	CCACCTCAAC-----	7117		
Qy	1326	CCTCATCTCTCATGTGCCCACTCCCATCGCTTCCCATTCGCCCAATTCGCGAATCCGGTTT	1385		
Db	7117	-----GGCAGAAATCCAGTTT	7131		
Qy	1386	GCC-CCTGCTCTCAACCCAGGAGCCCTGTTAGGCCCGATGTGAACCACTGACTTGA	1444		
Db	7132	CCCTTCTGCTATCAATCCAGGAACCCAGGCTGTGCTGGATGTTTT	7184		
Qy	1445	CCTCAGATCTCAGAGAAGCCAGTTTCATTTAATGTTCTGAGGGCGGCTTGAGATCC	1504		
Db	7184	-----TGGGGTTCAGAGAAATCAAGGCATCTCTGAGGGCGGCTTGAGATCG	7232		
Qy	1505	ACTGAGGGAGTGGTTTAGGCTCTGTGAGAGCAAGGTGAGATCTGAGGGAGGACTG	1564		
Db	7233	GCTGAGGGAGGCGGCCCAAGCTCTGTGGGAGCAAGGTGAGACTCTGAGGAAGGACTG	7292		
Qy	1565	AGGAGGACACACCCCAAGTAGTGGCCCAAAATGATCCAGTACCACCTCTGCTGCCAG	1624		
Db	7293	AGGAGGCGCCCAACCAAGTAGA-GGAACCAAAATATCCAGGCGAGCTCTCTGCTGCCAG	7351		
Qy	1625	CCCTGGACCCCGGCCAGGACAGATGCTCAGCTGGACCAACCCCGCTCCCGTCCCACT	1684		
Db	7352	TCCTGGACCCCGG---GGGAAGCTTCTCA---GGCTAGGCCATCCAGCTCCCACT	7404		
Qy	1685	GCCATTAAACCCACAGGCAATCTGTAGTATAGCTTATGTGACCCGGGCGAGGTTGGTC	1744		
Db	7405	GCCACTAAAGCTACAGGGGACTTAGAGTCAAGAGCTTGGTGTGCCCA-----	7453		
Qy	1745	AGGAGAGGAGGCGCCAGGATCAAGGTCAGCATCCCGCCGCAATTAGGTCAGGACCC	1804		
Db	7453	-----AGGAGGCGCCAGG-----TCTGCTGGCATCGGGGCTCAGGACT	7493		
Qy	1805	TGGAGGGAATGAGGTTCCCAACCCCACTGCTCTCTCTCATCTCCAGCCCAACCCCAAC	1864		
Db	7494	TGAGGGAATGAGGCGCTACACCCCAACCCCACTCCGATTC-----CAACAT	7543		
Qy	1865	TCATTTCCATACCTACCCCTACCCCAACCTTATCTTGTGAGAAATCCCTGCTGTCAA	1924		
Db	7544	GCCAGGCCCATCCCAACTCCGTTTTCAGAAATCCATTTT-----TCCCTGCACTCAA	7599		
Qy	1925	CCACGGAAGCCAGGGAATGGGCGGAGGACTCTGAGTCTTGCAGTCCCAATCCAGGCT	1984		
Db	7500	CCCGGGAAGACCTTGGGAATGGT---CAGGCACTCGGATCTTGACATCCACATCGAGGGC	7656		





Db	9812	CCATCAGCTTCACCTTGCTGGAGGCAACCAATGAGGGTTCCAGCAGCCAAAGAAGAGGAGG	9871
Qy	4140	GGCCAAGCACCTCTTGTTATCTCGAGTCCCTGTGTTCCGAGCAGTAATCACTAAGAAGTGG	4199
Db	9872	GGCCAAGCACCTCGCTGACGCGAGAGTCTTGTTCCGAGAAGCACTCAGTAACAAGGTGG	9931
Qy	4200	CTGATTTGGTTGGTTTGTCTCTCAAAATATCAGAGCCAGGAGCCAGTCAACAAGGCAG	4259
Db	9932	ATGAGTTGGCTCATTTCTGCTCCGAAGTATCGAGCCAAAGGAGTGCTCAACAAGGCAG	9991
Qy	4360	AAATGCTGGAGAGTGCATCAAAAATTACAAGCACTGTTTTCCTTGAGATCTTCGCGCAAG	4319
Db	9992	AAATGCTGGAGAGAGTCAATAAAAATTACAAGCGCTGCTTTCCTGTGATCTTCGCGCAAG	10051
Qy	4320	CCTCTGAGTCTCTCAGCTGCTGTTTGSCATTGACGTGAAGGAAGACAGACCCACCGGCC	4379
Db	10052	CCTCGAGTCCCTGAAGATGATCTTTGGCATTTAGTGAAGAAAGTGGACCCGCCAGCA	10111
Qy	4380	ACTCTATGTCTTGTGCACCTGCCTAGTGTCTCTCTATGATGGCTGCTGGGTGATAATC	4439
Db	10112	ACACCTACACCTTGTCACTGCCTGGGCCCTTCTCTATGATGGCTGCTGGGTATAATC	10171
Qy	4440	AGATCATGCCCAAGCAGGCTTCTGTATAATTGTCTTGGTCATGATGCAATGAGAGGCG	4499
Db	10172	AGATCTTTCCCAAGACAGGCTTCTGTATAATCTCTGGGCACAAATTCGAATGAGGCGC	10231
Qy	4500	GCCATGCTCTCGAGGAGAAATCTGGGAGGAGCTGAGTGTGATGAGAGTGATCATGGGA	4559
Db	10232	ACAGGCTCTGAGAGGAAATCTGGGAGGAGCTGGGTGTGATGGGTGTATGATGGGA	10291
Qy	4560	GGGAGCAGTGCCTATGTTGGGAGCCAGGAAGCTGCTCACCCAAGATTTGGTCAGGAAA	4619
Db	10292	GGGAGCACACTGTCTATGGGGAGCCAGGAACCTGCTCACCCAAGATTTGGTCAGGAAA	10351
Qy	4620	AGTACTTGGAGTA-CGGCAGGTGCCGACAGTATCCGCGACGCTATGAGTTCCTGTGG	4678
Db	10352	ACTACCTGGAGTACGGCAGGTACCGGCGAGTAATCTCTCGCGCTATGAGTTCCTGTGG	10411
Qy	4679	GTCCAAGGGCCCTCGCTCAACACAGCTATGTGAAGTCCCTTGACTATGTATCAAGTCA	4738
Db	10412	GTCCAAGGGCTCTGGCTGAACACAGCTATGTGAAGTCCCTGGAGCATGTGTCAGGGTCA	10471
Qy	4739	GTGCAAGAGTTTCGCTTTTCTTCCATCCCTGCTGAAGCAGCTTTGAGAGCAGGAAAG	4798
Db	10472	ATGCAAGAGTTGCGATTCGCTTACCCTACCCTGCTGAGCAGCTTTGTAGAGGAGAG	10531
Qy	4799	AGGAGTCTGAGCATGAGTTTCACGCCAAGGCCAGTGGAGGG- - -GCATGGGCCAGTG	4854
Db	10532	AGGAGTCTGAGCATGAGTTTCACGCCAAGGGCTGTGGGAGGAGGCCGCTGGGCCAGTG	10591
Qy	4855	CACCTTCCAGGGCCGCTGCAGCAGCTTCCCTGCCTGCTGTGACATGAGCCCATTCCT	4914
Db	10592	CATCTAACA- - -GCCCTGTGACAGAGCTTCCCTTGCCTGCTGTACATGAGGCCCATTCCT	10649
Qy	4915	CACATC- - - -TGAAAGAGCGGTCAAGTCTTCAGTAGTAGGTTTCTGTTCTATTGGGTGA	4970
Db	10650	CACATCTGTTGAAGAAATAGTCAAGTCTTCTAGTAGGGGTTCTATTGTTGGATGA	10709
Qy	4971	CTTGAGATTATCTTTGTTCTCTTTTGGAAATGTTCAAAATGTTTTTTTTTAAGGGATGG	5030
Db	10710	CTTGAGATTATCTCTGTTTCTTTTACAATTTGTAAGTGTG- - -TTCCTTTTAATGGATGG	10768
Qy	5031	TTGAATGAACTTCAGCATCCAAAGTTTATGAATGACAGCAGTCAACAGTCTGTGTATAT	5090
Db	10769	TTGAATTAACCTTCAGCATCCAAAGTTTATGAATCTAGTTAACGATATATGCTGTTAATAT	10828
Qy	5091	AGTTTAAGGGTAAAGAGTCTTGTGTTTATTTCAGATTTGGAAATCCATTCATTTTGTGAA	5150
Db	10829	AGTTTAGAGTAAAGAGTCTGTTTTTATTTCAGATTGGGAATCCGTTCTATTTTGTGAA	10888
Qy	5151	TTGGG- - -ATAATAACAGCAGTGGAAATAGTACTTTAGAAATGTGAAAAATGAGCAGTAA	5207
Db	10889	TTTGGGACATTAATAACAGCAGTGGAGTAAGTATTTAGAAAGTGTG- - -AATTCACCGTGAA	10945

Qy	5208	ATAGATGAGATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCCTGCTTATACCTCAG	5267
Db	10946	ATAGGTGAGAT-----AAATTAAGAGATCTTAATTCCTCCGCTTATGCCTCAG	10993
Qy	5268	TCTATTCTGTAAATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTCTTTTGAGA	5327
Db	10994	TCTATTCTGTAAATTTTAAAAATATATGCATACCTGGATTTCCTTGGCTC---GTGA	11050
Qy	5328	ATGTAAAGAGAATTAATCTGAATAAAGAATTCCTCTGTCTACCTGGCTCTTTTCTTCTC	5387
Db	11051	ATGTAAAGAGAATTAATCTGAATAAATAATTCCTCTGTAACTGGCTCATTTCTTCTC	11110
Qy	5388	CATGCACTGAGCATCTGCTTTTGGAAAGGCCCTGGGTAGTAGTGAGATGCTAAAGTAA	5447
Db	11111	TATGCACTGAGCATCTGCTCTGTGGAAGGCCAGAGATTAGTAGTGAGATACTAGGGTAA	11170
Qy	5448	GCCAGACTCATACCCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCAAGTACGAG	5507
Db	11171	GCCAGACACACACCTACCGATAGGGTATTAAAGACTTAGGAGCGCGTCAATAAATAAG	11230
Qy	5508	GTGCGAAGATGTCCTCAAGATCTAGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCT	5567
Db	11231	GTGACAAGATGTCCTCTAAGATGTAGGGGAAAAGT----AACGAGTGGGTATGGGGCT	11286
Qy	5568	CCGGGTGAGAGTGTGGAGTGTCAATCCCTGAGCTGGGGCATTTTGGGCTTTCGGGAAC	5627
Db	11287	CCAGGTGAGAGTGTGGGTGTAAATTCCTGTG-TGGGGCTTTTGGGCTTTCGGGAAC	11345
Qy	5628	TGCAGTTCCTTCTGGGGGAGCTGATTGATGATCTTTGGGTGATCC	5674
Db	11346	TGCATTTCCTCAGGGATCTGATTTCTAATGAGCTTGGTGGGTCC	11392
RESULT	6		
LOCUS	U82696		
DEFINITION	U82696	158898 bp DNA	PRI 01-JAN-1998
ACCESSION	U82696	Homo sapiens chromosome X map Xq28, complete sequence.	
NID	92735037		
VERSION	U82696.1	GI:2735037	

```

SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
AUTHORS     Kinschmann,B., Bauer,D., Drescher,B., Knop,A., Nordsieck,D.,
            Kioschke,P., Poustka,A., Rosenthal,A. and Gloeckner,G.
TITLE       Location of GABA receptor gamma-4 subunit (GABRG4) on Xq28
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 158998)
AUTHORS     Gloeckner,G., Rosenthal,A., Drescher,B., Schattevoy,R., Hinzmann,B.
            and Poustka,A.
TITLE       Direct Submission
JOURNAL     Submitted' (19-DEC-1996) Genome Analysis, Institut for Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
    source      1..158998
                . /organism="Homo sapiens"
                . /db_xref="taxon:9606"
                . /chromosome="X"
                . /map="Xq28"
    misc_feature 1..39138
                . /note="Region: cosmid IF064"
    repeat_region 1..106
                . /rpt_family="L1pAll"
                . /evidence=not_experimental
    repeat_region 213..282
                . /rpt_family="L1MB3"
                . /evidence=not_experimental
    exon          complement(1180..1284)
                . /note="GRAIL score = 43.000% comment = marginal"

```



```
/evidence-not_experimental
complement(1683. .1767)
/note="M2EF, score = 60%"
/evidence-not_experimental
complement(2062. 2120)
/note="FEXHB, score = 11.76%"
/evidence-not_experimental
complement(2637. .2710)
/rpt_family="L1ME2"
/evidence-not_experimental
complement(2859. 3028)
/rpt_family="L1ME3A"
/evidence-not_experimental
complement(2911. 2978)
/note="M2EF, score = 79.7%"
/evidence-not_experimental
4250. 4431
/note="GRAIL, score = 52.000%, comment = good shadow"
/evidence-not_experimental
complement(4639. 4866)
/rpt_family="L1MA2"
/evidence-not_experimental
complement(4893. 5203)
/rpt_family="L1MA2"
/evidence-not_experimental
complement(4974. 5117)
/note="GRAIL, score = 61.000%, comment = good"
/evidence-not_experimental
complement(5206. 5561)
/rpt_family="L1MA2"
/evidence-not_experimental
complement(5562. 5602)
/rpt_family="L1MA9"
/evidence-not_experimental
complement(5619. 5655)
/rpt_family="L1MA5"
/evidence-not_experimental
complement(6084. 6254)
/rpt_family="MLT1D"
/evidence-not_experimental
complement(6262. 6322)
/rpt_family="MLT1D"
/evidence-not_experimental
complement(6384. 6549)
/rpt_family="MLT1D"
/evidence-not_experimental
6664. 6771
/note="homology = 89.80%, score = 82, counts = 27"
/rpt_type=tandem
/rpt_unit=ccctt
/evidence-not_experimental
6807. 6884
/note="FEXHB, score = 10.83%"
/evidence-not_experimental
7337. 7454
/rpt_family="L1"
/evidence-not_experimental
7483. 7533
/note="FEXHB, score = 12.95%"
/evidence-not_experimental
7950. 8010
/rpt_family="MSTA"
/evidence-not_experimental
8147. 8238
/rpt_family="MSTA"
/evidence-not_experimental
8245. 8286
/rpt_family="MSTA"
/evidence-not_experimental
8287. 8361
/rpt_family="MSTC"
/evidence-not_experimental
8878. 8921
```

```
/note="homology = 97.70%, score = 20, counts = 2"
/rpt_type=tandem
/rpt_unit=gattttacctgtcaagtcta
/evidence-not_experimental
complement(9471. 9526)
/note="M2EF, score = 85.6%"
/evidence-not_experimental
complement(9474. 9529)
/note="FEXHB, score = 12.02%"
/evidence-not_experimental
9618. 9660
/rpt_family="MER3"
/evidence-not_experimental
10066. 10099
/note="GRAIL, score = 61.000%, comment = good"
/evidence-not_experimental
11110. 11190
/note="GRAIL, score = 76.000%, comment = excellent"
/evidence-not_experimental
11257. 11275
/note="FEXHB, score = 19.61%"
/evidence-not_experimental
11360. 11419
/note="homology = 91.70%, score = 20, counts = 2"
/rpt_type=tandem
/rpt_unit=agggaataagtagaactatcatcaggc
/evidence-not_experimental
11613. 11637
/note="xpond exon prediction, score = 68% (0%)"
/evidence-not_experimental
complement(12630. 12662)
/note="FEXHB, score = 13.07%"
/evidence-not_experimental
13530. 13600
/rpt_family="MLT2B2"
/evidence-not_experimental
13607. 13760
/rpt_family="MLT2B2"
/evidence-not_experimental
13820. 13947
/note="GRAIL, score = 88.000%, comment = excellent"
/evidence-not_experimental
13924. 13977
/rpt_family="MLT2D"
/evidence-not_experimental
13978. 14043
/note="homology = 75.80%, score = 23, counts = 6"
/rpt_type=tandem
/rpt_unit=ctctctccttt
/evidence-not_experimental
14054. 14087
/note="homology = 94.10%, score = 28, counts = 17"
/rpt_type=tandem
/rpt_unit=tg
/evidence-not_experimental
complement(15048. 15137)
/note="FEXHB, score = 12.14%"
/evidence-not_experimental
complement(15049. 15138)
/note="M2EF, score = 83.8%"
/evidence-not_experimental
15192. 157359
/note="Region: Cosmid IG0799"
complement(15720. 15884)
/note="GRAIL, score = 53.000%, comment = good"
/evidence-not_experimental
16505. 16631
/rpt_family="L1MA9"
/evidence-not_experimental
16659. 16746
/rpt_family="MER4A"
/evidence-not_experimental
16798. 16851
```

/rpt\_family="MER4A"  
/evidence-not\_experimental  
complement(16851..16910)  
/rpt\_family="THE1B"  
/evidence-not\_experimental  
complement(16920..16972)  
/rpt\_family="THE1B"  
/evidence-not\_experimental  
complement(16931..16963)  
/note="FXHB, score = 12.61"  
/evidence-not\_experimental  
complement(17023..17178)  
/rpt\_family="THE1B"  
/evidence-not\_experimental  
complement(17252..17279)  
/rpt\_family="THE1B"  
/evidence-not\_experimental

Query Match 46.8%; Score 2653.2; DB 11; Length 158898;  
Best Local Similarity 73.6%; Pred. No. 0;  
Matches 4230; Conservative 0; Mismatches 1103; Indels 415; Gaps 50;

QY 21 CTTCCCTACCAACCCCAATCCCTTACGCCACCCATCCAAACATCTTCAGCGTC 80  
DB 57500 CCAGCACCCCTATCTCCCAACCCCACTACCTTATGTCTTATCTCCCAACCCCAAC 57559  
QY 81 ACCCCAGCCCAAGCAGGAGGAATCCGGTTCACACCCCTCTCTCAACCCAGGAGGC-C 139  
DB 57560 ACCATATCCCATCCAGGTGTAATCGATTCCTCTCTCAACCCAGGAGGC 57619  
QY 140 CAGGTGCCAGATGTGACGCCATGACTTACGATTTAGTGGTTAGAGAGGAGGATTT 199  
DB 57620 CAGGTCTGGATGTGATGCCAGTACTTGTGATGGGTTAGAGAGCGCTAGCTTC 57679  
QY 200 TCGTCTGAGGGGGCTTGCAGTCTGGTGGAGGAGCGGCCCA-GCTGTGAAGGAGG 258  
DB 57680 TCAGTCTGAGGAGCGCTTGGGATGAGGAGGAGCGGCCGCTGTGTGAGGTGG 57739  
QY 259 CAAGGTGACATGTGAGGAGGACT---GAGGAGCCCACTTACCCAGATAGAGAGCCCA 315  
DB 57740 CATAGTGAAGAGCTGAGGAGAGTGGGAGGCCCTCTCCACCCAGATAGAGACCCCA 57799  
QY 316 AATAATCC-----CTTCATGCCAGTCTTGAGACCATCTGTGTGGTACTTCTCAGGC 366  
DB 57800 AATAATCCAGCCCTCTCTGCTTCCAGTCTGGGACACCGTGGGCGGACTTCTCAGTC 57859  
QY 367 TGGGCCACCCAGCCCTCTGCTTAAACCACTGGGAGCTC-GAAGTCAGAGCTCCG 425  
DB 57860 TGGAGCCCAACCCCACTGCGCTGGAAGCCGAGGAGGACTATGGATCAGAGCTTGG 57919  
QY 426 TGTGATCAGGAAAGGCTGCTTAGGAGAGGCGAGCGTCCAGGCTCTGCCAGACATATGC 485  
DB 57920 TGTGATCAGTGCAGGACTGTGGGGT-----AGGCTCTGCCAGGATCAAGC 57967  
QY 486 TCAGATTTCAAGAGGGGTGAGGTCTCTAGACCCCACTCCGCTGAGACCCCAACCCCA 545  
DB 57968 TCAGGACCTTAGGAGGGGTGAGTGTCCCA---CCCCCATTCCTATCCCTCAGCT 58024  
QY 546 CTCATGCTCACTCCGTCAGCCCAACCCCTCTTCATTTGATTCATCCACCCCAACCCCA 605  
DB 58025 TTCCATCTGCATCTCC-----TACCCATCTGTACCCC 58058  
QY 606 CATCCCAACCCCACTCTCAAGCCCTGATGCCATCCGCCAGCCATTCACCCCTCACCC 665  
DB 58059 CATTCACCACTGTGCCCTATCTCTCCCAACCCCAACCCAGCCCTATACCCCTCC 58118  
QY 666 CCACCCCAACCCCAAGCCCACTCCCAACCCCAACCCAGGAGGATCGG-GTTCGCCGAG 724  
DB 58119 CCACCCCTATCTATCCCATCAGTGCAGCATCCGTTCCACCCCTCTCTTCAATCCAG 58178  
QY 725 GAAACATCCGGGTGCGGATGTGAGCCCACTGACCTGCTTCCATCTCCACCCCAACCCCA 784  
DB 58179 GCAAGCCCTGGGTGGCGGATGTGATGCCACTGACTTGTGAATTGAGGTTAGAGAGAG 58238

QY 785 CGAGGTTTCAATCTTGAGGACGCGGTAGAGTTCGGCCGAAGAACCTGACCCAGGCTCT 844  
DB 58239 TGAGTTTCTGGGTCTGAAGGTTGGC-TTGAGATCGCAGAGGGAAGGTGCCAGGCTTT 58297  
QY 845 GTGAGAGGCAAGGTGAGAGGCTGAGGAGGACTGAGGACCCGCCCACTCCAATAGAGA 904  
DB 58298 GTGAAGAGGCAAGGTGAGACTCTGAGGAGGATTGAGAAACCCCTATCCCTGATAGAG 58357  
QY 905 GCCCAAAATATTCAGCCCGCCCTTCTGCCAGCCCTGCCCCACCGCGGGAAGAGCTC 964  
DB 58358 GTCCAGCCCTGAGACTACCC-----CGCGAGGCTGACTTC 58393  
QY 965 TCAGCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACAGGTTCTTCT 1024  
DB 58394 TCAGACTGGGCTGCTCCCAACCTCCGCCCC---TTCCACAGCGCTTGTTAACCCA 58448  
QY 1025 CCCAAGCTCTGGAATCAGAGGTTGCTGACAGGCGAGGACTGGTTAGAGAGGCGCAG 1084  
DB 58449 CAGGGGACTCTGGAGTCAGAGTTGGTGTGATCAGGGAAGGCTGGTTAGGAGAGGCAT 58508  
QY 1085 GGCACAGGCTCTGCCAGGCAATCAAGATCAGCACCCCAAGAGGAGGCTGTGGCCCA 1144  
DB 58509 GGCACAGGCTCTGCCAGGCAATCAAGATCAGAACCC-TGAGAGGGAATGAGGTCCCA 58567  
QY 1145 GACTGCACTCCAATCCCACTCCCAACCCCAATTCGCATTTCCCAACCCCAACCCCA 1204  
DB 58568 GATCCTAGTCTAACCCCACTCCCA-------- 58595  
QY 1205 TCTCTCAGTACACCTCCACCCCACTCCCTACTCTACTCTCCTCACTGACCAACCC 1264  
DB 58595 -----AATCGCTGCCATTTCTGCTGCTTCCATTTCCATTTCTTCCCT 58636  
QY 1265 TCCAGCCCAAGCAGCCAGCCCAACCCCTTCTGCCACCTTCCCTCACTGCCCCCAACCCCA 1324  
DB 58637 TCCACCTTCCCA-------- 58650  
QY 1325 CCCTCATCTCTCATGTGCCCCCACTCCCATCGCTCCCTCCCTTCTGGCAGAATCCGTT 1384  
DB 58650 -----GGCAGAATCCAGTT 58663  
QY 1385 TGCC-CCTGCTCTACCCAGGAGGAGCCCTGGTAGGCCCATGTGAACCACTGACTTGA 1443  
DB 58664 CCCTTCTGCTATCAATCCAGGAAACCCCAAGCTTGGTCTGGGATGTTTT----- 58717  
QY 1444 ACCTCAGATCTGAGAGAGCCAGGTTTCATTTAATGGTTCTGAGGGCGGCTTGAGATC 1503  
DB 58717 -----TGGGGTTCAGAGATCAAGGCAATAGTCTGAGGGCCAGTTGAGATC 58764  
QY 1504 CACTGAGGAGTGGTTTATAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGAGGACT 1563  
DB 58765 GGCTGAGGAGCGGCCCAAGCTCTGTGCGAGGCAAGGTGAGACTCTGAGGAAGGACT 58824  
QY 1564 GAGGAGCAGACACCCAGGTAGATGCCCAAAATGATCCAGTACCACTCCCTGCTGCCA 1623  
DB 58825 GAGGAGGCCCAACCCCAAGATAGA-GGAACCAAAATCAATCCAGCGAGCTCTGCTGCCA 58883  
QY 1624 GCCTGAGCCACCCGCGCAGGAGCAGATGCTCAGCTGGACCAACCCCGCTCCGTCAC 1683  
DB 58884 GTCTGAGCAGCCCGG---GGGAAGACTTCTCA---GGTAGGCCATCCAGCTCCAC 58936  
QY 1684 TGCCACTTAACCCACAGGCAATCTGTAGTCAATAGCTTATGTACCGGGGAGGTTGGT 1743  
DB 58937 TGCCACTAAAGCTACAGGGGACTTAGAGTCAAGAGCTTGTGTGCCA----- 58986  
QY 1744 CAGGAGGAGGCGGCCAGGCATCAAGGTCCAGCATCCCGCCCGGCTATAGGTCAGGAC 1803  
DB 58986 -----AGGCAGGCGCCAGGC-----TCTGCTGGCATCGGGTCAGGAC 59025  
QY 1804 CTGGGAGGGAATGAGGTTTCCCAACCCCAACCTGCTCTCTCATCTCCACCCCAACCCCA 1863  
DB 59026 TTGAGAGGGAATGAGGCGCTACACCCCAACCCCACTCCGATTC-----CAACA 59075

Qy	1864	CTCACATTCCTCCATACCTACCCCTTACCCCCAACCTCATCTTTGTGCAGAAATCCCTGCTGTCA	1923
Db	59076	TGCCAGCCCCATCCCCAACTCCGTTTGTGAGAATCCATTTTT---TCCCTCGACTCA	59131
Qy	1924	ACCCACGGACACACGGGATGGCGGCAGGCACTCGGATCTTGACCTCCCCATCCAGGG	1983
Db	59132	ACCCCGGAAGACCTGGGAATGGT---CAGGCATCTCGATCTTGACATCCACATCGAGGG	59188
Qy	1984	TCTGATGGAGGAAGG-----GCTTGAACAGGGCTCAGGGGACAGAGGGAG---	2033
Db	59189	CTGAAGAGGGAGAGGTTTGGTATCATGACGAGCCTCAGGTAGCAGGGGAGGACC	59248
Qy	2033	--GGCCCTACTGCGAGATGAGGGAGGCTCAGAGGACCCAGCACCTTAGACACCG-----	2087
Db	59249	CTGGCCCTCTGGGAGATGAGGAAGGCTCAGGAGACCCAGCACCCCAAGCAGGGAGCC	59308
Qy	2087	---CACCCCTGCTGAGACTGAG--GCTGCCACTTCTGGCCCTCAAGAATCAGAACCATG	2140
Db	59309	CACCCACCCCGTCTGAGAATGAGTGCTCTCTTTAGCCTCAGGAATCCAGGGATG	59366
Qy	2141	GGGACTCAGATTGCATGGGGTGGA---CCAGGCGCTGCAAGGCTTACGGGGAGGAAGAG	2198
Db	59369	GCACTCAGGTCACGAGGGGTGGTTCCAAGCCCTCCAGGATCAAGGAAGAGAC	59428
Qy	2199	GAGGAGGACTCAGGGACCTTGGAAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTC	2258
Db	59429	GAGGAGGATTCAGGGGCGCTTGCATTCCAGATCAGTGGAGACCTGGGCCCTGGGAGGTC	59488
Qy	2259	CAGGGCAGGTGGCCACATATGGGCCCATATTTCTTGATCTTTGAGT---GACAGGAC	2314
Db	59489	CTGGCAAGGTAGGCACCTGTAGTCTCATCTTCTGCTACTCTCGAGGTCAAGAGAGGAG	59548
Qy	2315	AGGCTCTGGTCTCAGAAGTGGGGCGCTCAGGTCAACAGAGGGAGGAGTTCACAGATCCAT	2374
Db	59549	AGGCTATGTTCTCAGGGGTGGTACTTCAGTCCGACAGGGAGAGAGTCCAGGATCTAC	59608
Qy	2375	ATGSCCAAGATGTGCC--CCCTTCATGAGGACTGGGATATCCCCGCTCAGAAGAAGG	2433
Db	59609	AGGACCAAGGTGTGCCACACTTCACGAGGAATGGGATACCTGTGGCTCAGAAGAAGCG	59668
Qy	2434	GACTTCCACAGTCTGGCTGTCCCTTTTAGTACTCTAGGGGACCATCAGGATGG	2493
Db	59669	GACCCACAGAGTCTGGCTGTCCCTGTCTTAGCTCAGGGGGACCATCAGAGGAGGATGG	59728
Qy	2494	CGGTATGTTCAATCTCATTGTACACAGCAGGAAGTGGGGGGCGCTCAGGGAGATG	2553
Db	59729	CCCTATGTCCAAATTCATTGTGTTCCAGCGCAGAGTGGGGAACCTTCAGGGAGATG	59788
Qy	2554	GGGTCTTGGGTAAAGGGGATGCTACTCATGTCAGGGAATTTGGGGTTGAGGAAGCA	2613
Db	59789	AGGTTTTGGAGTAAAGGGCAATGTTGCTCATCTCAGGGGTTGGGGTTGAGGAAGGG	59848
Qy	2614	CAGCGCTGGCAGGAATAAGATCAGTGTAGACACACAAAGCTATTGGAATCCACACCCA	2673
Db	59849	CAGGCCCTGTGACAGGACAAACATGAGTACCCACAGAGGCGCATCAGAACCTCACCCCA	59908
Qy	2674	GAACCAAGGGTCAGGCCCTGGACACTTCACCCAGGATGTG-----GCTTCTTTTTCT	2725
Db	59909	GAACCAAGGGTCAGGCCCTGGGACCCACACAGGGGTGACAGGATGTGCTCTCTCTC	59968
Qy	2726	ACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCTCATTTCTCAGAGGTTGACTCAGTCA	2785
Db	59969	ATTTCTGATCCAGATCTCAGTGTAGGTGAGGACCTTGTCTCTCAGAGGTTGACTCAGTCA	60028
Qy	2786	ACGTAGGGACCCCATCTGGTCTTAAAGACAGAGCGGTCCCGAGGATCTGCCATGCGTTGG	2845
Db	60029	CCACAGGGACCCCATCTGGTCTTACAGACAGTGTGTCCCGAGGATCTGCCAAGAGTCTCTG	60088
Qy	2846	GTGAGGAACATGAGGGAGGACTGAGGTTACCCAGGACCCAGAACACTGAGGGACACTGCA	2905
Db	60089	GTGAGGAATGTGAGGGAGGATTTGAGGTTACCAAGGCCAGAACCGCAGATGATGACCCCA	60148
Qy	2906	CAGAAATCAGCCCTGCCCTGCTGTACCCCGACAGAGCATGGGCTGGGCGGTCTGCCGAG	2965

Db	30149	CAGAAATCAGCCCTGCCTCCTTGTCTCACCACAGAGAGCATGGGCTTGCTTTCTCTGCTGAG	60208
Qy	2956	GTCCCTCCGTTATCCTCGGATCATTGATGTACAGGACGGGAGGCCCTTGGTCTCAGAAGG	3025
Db	50209	GTCCCTCTCTTATCCTCGGATCACTGGTGTCCAGAGTGGAGAGCCCTTGGTCTCAGGGGG	60268
Qy	3026	CTGGCCTCAGTCACTAGAGGAGCGTCCCAGGCCCTGCCAGGAGTCAAGGTGAGGACCA	3085
Db	60269	CTGCACCCAGCTCAGTAGAGGAGGGTCCCAGGCTCTGCCAGGAGTTCAGGTGAGGACCA	60328
Qy	3086	AGCGGGACCTCACCCAGGACACATTAATTCCAATGAATTTTGATATCTCTTGTCTGCCT	3145
Db	60329	AGCAGGCTCCGCATCCACGACACATGGGTTTCCAATGAATTTTCGACATCTTTTGTCTGCT	60388
Qy	3146	TC-CCCAAGGACCTAGGCACGCTGTGGCCAGATGTTTGPCCCTCTCTGTCCTTCCATCCCT	3204
Db	60389	TCTTCGGNAGACCTTAGGCACAGTGGCCAGATGTGGGGTTTCTTAGTTCCT--GTTCCC	60445
Qy	3205	TATCATGGATGTAACTCTTGATTGGATTTCTTCAGACCAGCAAAAGGCGCAGGATCCAGG	3264
Db	60446	TCTCAGGCATCTGAGCTCTTGATCTGATTTCTCAGGCCAGCAAAAGAGTGGGATCCAGG	60505
Qy	3255	CCCTGCCAGGAAAATATAGGGCCCTCGTGAACAACAGAGGGGGTCACTCACATGTCATGA	3324
Db	60506	CCCTGCCTTGGAGAAATGTAGGGCCCTGAGTGAACACAGTAGTGGGATCATCCACATCCATGA	60565
Qy	3325	GAGTGGGATCTCACAGATCCAGCCACCCCTCTGTTAGCACGTGAGAACCCAGGGCTGT	3384
Db	60566	GAGTGGGACCTCACAGAGTCCAGGCTACCCCTTTGATGGCACGTAGGGACCGGGGCTGT	60625
Qy	3385	GCTTCGGCTTCGACCTCTAGGGCCCTGGATTCTCTCTCTGGAGTCCAGGAACACAGG	3444
Db	60626	GCTTACAGTCTGCACCTAAGGGCCATGATTCTCTCTTAGGAGTCCAGGAAACAGG	60685
Qy	3445	CAGTAGGGCTTGGTCTGAGACAGATACCTCAGGTACAGAGCAGAGATGCACAGGGTG	3504
Db	60686	CAGTAGGGCTTGGTCTGAGACAGTGTCTCAGGTTACAGAGCAGAGATGCACAGGGCTG	60745
Qy	3505	TGCCAGCAGTCAATGTTTGCCCTCAATGCACACCAAGGCCCCACCTGCCACAGACACA	3564
Db	60746	TGCCAGCAGTGAATGTTTGCCCTGAATGCACCAAGGGCCCCACCTGCCACAAGACACA	60805
Qy	3565	TAGGACTCCACAGAGTCTGGCCTCACCTCCTACTGTCACTGTAGTGAATCGACCTCTG	3624
Db	60806	TAGNCTCCAAAGAGTCTGGCCTCACCTCCTACCATCAATCCTGCAGATCGACCTCTG	60865
Qy	3625	CTGCGCGCTGTACCTGA-GTACCCCTCTCACTTCTCTCTTTCAGGTTTCAGGGGACAGG	3683
Db	60866	CTGCGCGCTATACCTCAGGTGCTCTCTCACTTCTCTCTTTCAGTTCGTGAGCAGACAGG	60925
Qy	3684	CCAAACCCAGGACAGGATTCCTTGAGGCCACAGAGGAGCACCAGGAGAAGATCTGTA	3743
Db	60926	CCAAACCG-GAGGACAGGATTCCTTGAGGCCACAGAGGAGCACCAGGAGAAGATCTGTA	60984
Qy	3744	AGTAGGCTTTGTTAGAGTCTCCAAGTTTCAGTTCTCAGCTGAGGCCCTCTCACACACTCC	3803
Db	60985	AGTAGGCTTTGTTAGAGCTCTAAGATTGGTTCCTCAGTCTCCTCACATGCTCT	61044
Qy	3804	CTCTCTCCCCAGGCTGTGGGTCTTCAATPGCCCAAGTCTCTGCCACACTCTGCTGCTGTG	3863
Db	61045	CTCTCTCGTAGGCTGTGGGTGCCCATTTGCCAGCTTTTGCCTGCATCTTGCCTGTCTG	61104
Qy	3864	CCCTGACGAGAGTCACTATGCTCTTGTGACAGAGGAGTCTGCATGCAAGGCTGAGGAG	3923
Db	61105	CCCTGACGAGAGTCACTATGCTCTTGTGACGAGAAGAGTCAGCACTGCAAGGCTGAGGAG	61164
Qy	3924	CCCTTAGGCCCCAACAGAGGCTTGGGCTGTGTGTGTGTCAGGCTGCCAC-----	3976
Db	61165	CGTTTAGGCCCAAGAGAGGCCCCGTGGGCTTGGTGGTGTGACAGGCTCTCTACTGAGG	61224
Qy	3976	-----CTCCCTCTCCTCTCCTCTGCTGTCTGGGCAACCTTGGAGGAGGTGC	4019



```
/sex="female"  
/cell_type="lymphocyte"  
/tissue_type="blood"  
/dev_stage="adult"  
2777..2848  
/number=2  
2924..4491  
/number=3  
2989..3942  
/gene="MAGE4b"  
2989..3942  
CDS  
/gene="MAGE4b"  
/codon_start=1  
/product="MAGE-4b antigen"  
/protein_id="AAA68872.1"  
/db_xref="pid:g533517"  
/db_xref="gi:533517"
```

```
/translation="MSSEQKSHCKPBEVEAQEALGLVGAQPTTEQBAAYSSSS  
PLVPTUEEVPAASAGPPQPOGASALPTTISFCWRQNEGSSSEEEGPPSTPDA  
ESLREALSNNKVDLSAHLFLRKYRAKELVTRKELRVKRYKRFVIFGKASLSK  
MIFGIDVKEVDPTNTVLTCLGLSYDLGNQIFPKTGLLIIVLGTIAEGDSAS  
EEELWEELGVNGVDGREHTVYGEPRKLLTQDWQENLYEVROVPGSPNARYEFLWGP  
RALAETSVKVLEHVVRVNRVRIAYPSLREALLLEEEGV"
```

```
polyA_signal 4472..4477  
BASE COUNT 1134 a 1235 c 1433 g 1093 t  
ORIGIN
```

Query Match 44.3%; Score 2513.6; DB 9; Length 4895;

Best Local Similarity 79.9%; Pred. No. 0;  
Matches 3415; Conservative 0; Mismatches 694; Indels 165; Gaps 32;

```
QY 1478 ATGGTCTCAGGGCGGGCTTGAGATCCACTGAGGGAGTGGTTTATAGGCTCTGTGAGGAG 1537  
DB 607 ATAGTCTCAGGGCGCAGTTGAGATCGCTGAGGGAGCGGGCCCAAGCTCTGTGGCGAG 666  
QY 1538 GCAAGGTGAGATGCTGAGGGAGGACTGAGGAGGCACACACCCAGGTAGATGGCCCCAAA 1597  
DB 667 GCAAGGTGAGATCTGAGGAGGAGCTGAGGAGGCGGCCACCCCAAGATAGA-GGAACCCAA 725  
QY 1598 ATGATCCAGTACCACCCCTGCTGCCAGCCCTGGACCACCCCGCCAGGACAGATGCTCAG 1657  
DB 726 ATAATCCAGCCACAGTCTGCTGCCAGTCTTGGACACACCCG- -GGGAAGACTTCTCA- 782  
QY 1658 CTGGACACCCCGCTCCGCTCCCACTGCACTTAACCCACAGGCAATCTAGTCTATA 1717  
DB 782 - -GGCTAGGCCATCCAGCTCCCACTGCCACTAAAGCTACAGGGAGCTCTAGAGTCA- - 837  
QY 1718 GCTTATGTACCGGGCAGGGTGTGTCAGGAGAGGCAGGGCCCGCCAGGCATCAAGGTCCAGC 1777  
DB 837 - - - - -AGAGCTTGGTGTGTCGCCAAGGCGAGGGCCAGG 866  
QY 1778 ATCCCGCCGGCATTAGGTCAGGACCTGGGAGGAACTGAGGGTTCCCCACCCACACCT 1837  
DB 867 CTCCTCCCTGGCATCGGGCTCAGGAGCTTGAGAGGAACTGAGGGGCGCTACACCCCCACCC 926  
QY 1838 GTCTCCTCATCTCCACCGCACCCCACTCACATTCCTACCTACCCCTACCCCAACCC 1897  
DB 927 CATCCGCATC- - - - -CAACATGCCAGCCCACTCCCACTCCGTTTGCAGNA 976  
QY 1898 TCATCTTGTAGAAATCCCTGTGTCAACCCACGGAAGCCACGGAATGCGGCCACAGCAC 1957  
DB 977 TCCATTTT- - - - -TCCCTCTGCAGTCAACCCCGGAAGACCTCGGAATGT- - -CAGGCAC 1029  
QY 1958 TCGGATCTTGACCTCCCACTCAGGGTCTGATGGAGGGAAGG- - - - -GCTTGAACAG 2010  
DB 1030 TCGGATCTTGACATCCATCAGCGGCTGAAGAGGGGAGAGATTTGTTATCATGAGCAG 1089  
QY 2011 GGCCTCAGGGAGCAGAGGAG- - - - -GGCCCTACTGCGAGATGAGGAGGCGCTCAGAG 2064  
DB 1090 AGCCTCAGGGTAGCAGAGGAGGAGCCCTGGCCCTCCTGGGAGATGAGGAGGCGCTCAGAG 1149  
QY 2065 GACCCAGCACCC- - - - -TAGGACACCGCACCCCTGCTGTGAGACTGAG- -GCTGCCA 2113
```

```
DB 1150 GACCCAGCACCCCAAGGAGGAGGACCCACCCACCCCTGTCTGAGAATGAGGTGCTCCT 1209  
QY 2114 CTTCTGGCCTCAAGAAATCAGAACGATGGGACTCAGATTGATGGGGTGGGA--CCCAAG 2171  
DB 1210 CTTTATGCTCAGGAATCAAGGGATGCAACTCAGGTCAGCAGAGGGGTGGTTCACAG 1269  
QY 2172 GCCTCAAGGGCTTACGCGGAGGAAGAGGAGGAGGACTCAGGGGACCTTGGGAATCCAGAT 2231  
DB 1270 CCCTTCCAGGATCAAGGAAAGGAGACGAGGAGGATTCAGGGGCGCTTGCATTCAGAT 1329  
QY 2232 CAGTGTGAGACCTCGCCCTGAGAGGTTCAGGCGACGCTGGCCACATATGGCCCATATTC 2291  
DB 1330 CAGTGGAGACCTGGGCCCTGGGAGTCTGGGCAAGGTAGCCACCTGTAGCTCATCTTC 1389  
QY 2292 CTGCATCTTTCAGGT- - -GACAGGACAGAGCTGTGGTCTGAGAAAGTGGGGCTCAGGTC 2347  
DB 1390 CTGCATCTTCAGGTCACAGAGAGGAGGGCTATGTTCTGAGGGGTGGTACTTCAGGTC 1449  
QY 2348 AACAGAGGAGGAGTTCAGGATCCATATGGCCCAAGATGTGCC--CCCTTCATGAGACT 2406  
DB 1450 CGCAGAGGAGGAGTCCCGAGGATCTACAGGACCCCAAGGTGTGCCACACTTCACGAGGAAT 1509  
QY 2407 GGGGATATCCCGGCTCAGAAAGAGGAGCTCCACACAGTCTGGCTGTCCCTTTTAGTA 2466  
DB 1510 GGGGATACCTGTGGCTCAGAAAGAGCGGACCCACAGAGTCTGGCTGTCCCTGTCTTTA 1569  
QY 2467 GCTCTAGGGGACACAGATCAGGGATGGCGTATGTTCCATCTCACTTGTACACAGGCA 2526  
DB 1570 GCTCAGGGGACACAGAGGAGGTGGCCCTATGTGCCAATTTCACTTGTTCACAGGCA 1629  
QY 2527 GGAAGTTGGGGGCGCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGATGCTACTCAT 2586  
DB 1630 GGAAGTTGGGGAACCTTCAGGGAGATGAGGTTTGGAGTAAGGGGCAATGTTGCTCAT 1689  
QY 2587 GTCAGGGAATTTGGGGTGTGAGGAACACAGGCGCTGGCAGGAATAAAGATAGTGAGACA 2646  
DB 1690 CTCAGGGGTTTGGGGTGTGAGGAAGGCGGCCCTGTGAGGAGCAACATAGT-ACCCA 1748  
QY 2647 GACAAGGTATTTGAATCCACACCCAGAACCAAGGGGTGAGCCCTGGACACCTCACCC 2706  
DB 1749 CAGGAGGCACTCAGAAACCTCACCCCAAGCAAGGGGTGAGCCCTGGGACCCACAC 1808  
QY 2707 AG- - - - -GATGTGGCTTCTTTTTCATCTCTGTTTTCAGATCTGGGACGAGTGAAGAC 2758  
DB 1809 AGGGGTGACAGGATGTGGCTCTCTCTCTCTCTGATTCAGATCTCAGTGAAGTGAAGAC 1868  
QY 2759 CTCATTTTCAGAGGTGACTCAGGTCAACGTAGGAGCCCCCATCTGGTCTAAAGACAGAG 2818  
DB 1869 CTTGTTCTCAGAGGTGACTCAGGTCAACAGGAGGCCCTCTGGTCTACAGACACAG 1928  
QY 2819 CGGTCCAGGATCTGCCATGCGTTCCGGTGAAGACATCAGGGAGGACTGAGGGGTACCC 2878  
DB 1929 TGGTCCCAAGGATCTCCCAAGAGTCTCTGTGAGGAATGTGAGGGAGGATTTGAGGGTACCAC 1988  
QY 2879 AGGACCAAGCACTCAGGGAGGACTGCAGAAATCAGCCCTCGCCCTGCTGTCAACCCAG 2938  
DB 1989 AGGGCCAGNACGCAATGATGACCCCAAGAAATCAGCCCTGCTCTGTGTGACCCAG 2048  
QY 2939 AGAGCATGGGTGGGCCCTGTGCCAGGTCTCTCGTTATCTTATCTGGGATCATTTGATGTAG 2998  
DB 2049 AGAGCATGGGTGGCTTGTCTGAGGGGTGTCACCCAGGTCTGATCTGATCTGATCTGATCT 2108  
QY 2999 GGACGGGAGGCGCTTGGTCTGAGAAAGGTGCGCTCAGGTCTAGGAGGAGGCTGCCAGG 3058  
DB 2109 GGAGGGGAGGCGCTTGGTCTGAGGGGTGTCACCCAGGTCTGATAGGGAGGCTGCCAGG 2168  
QY 3059 CCCTGCCAGGAGTCAAGGTGAGGACCAAGCGGGACCTCACCCAGGACACATTAATTTCA 3118  
DB 2169 CTCCTGCCAGGAGTTCAGGTGAGGACCAAGAGGCTCCCGCATCCAGGACACATGGGTCCA 2228  
QY 3119 ATGAATTTTATATCTTGTGCTGCTTTC-CCCAAGGACCTAGGACGCTGTGGCCAGATG 3177
```

[illegible]

Qy	4233	GAGCCAGGGACCGCAGTCAAAAGGCAGAATAAAGCTGTGAGAGTGTCATCAAAAATTACAAGC	4299
Db	3365	GAGCAAAGAGCTGGTCACAAAGGCAGAATAAAGCTGTGAGAGTGTCATCAAAAATTACAAGC	3424
Qy	4293	ACTGTTTCTTGAGATCTTCGGGCAAGCCCTCTGAGTCTCTTGACGTGCTCTTGGCATTTG	4352
Db	3425	CCTGCTTCTTGTAICTTCGGCAAGCCCTCGAGTCCCTGAAGTAGATCTTTGGCATTTG	3484
Qy	4353	ACGTGAAGAGACGAGACCCCACCAGGCCACTCTCATATGTCCTTGTCACCTGCCTAGGTCTCT	4412
Db	3485	ACGTGAAGGAAGTGGACCCCACCAGCAACACCTACACCCCTTGTCTACCTGCCTGGGCCCTT	3544
Qy	4413	CCTATATGGCCTGCTCGGTGTATAATCAGATCATGCCCAAGACAGGCTTCTTGATAATTG	4472
Db	3545	CCTATATGGCCTGCTCGGTGTAATAATCAGATCTTTCCCCAAGACAGGCCCTCTTGATAATCG	3604
Qy	4473	TCCTGGTCAATGTCGAATGGAGGGCGGCCACTGCTCTCGAGGAGGAATACTGGGAGGAGC	4532
Db	3605	TCCTGGGCACAATTGCAATTGGAGGGCGACAGCGCCTCTGAGGAGGAATACTGGGAGGAGC	3664
Qy	4533	TGAGTGTATGGAGGTGTATGATGGAGGGAGCACAGTGCCTATGGGAGGCCACGAGAAC	4592
Db	3665	TGGTGTGTATGGGGGTGTATGATGGAGGGAGCACACTGTCTATGGGAGGCCACGAGAAC	3724
Qy	4593	TGCTCACCCAAGATTTGGTGCAGGAAAACTACTCGAGTA-CGGCAGGTGCCGACACAGTG	4651
Db	3725	TGCTCACCCAAGATTGGGTGCAGGAAAACTACTCGAGTACGGCAGGTACCCGGCAGTA	3784
Qy	4652	ATCCCGCACGCTATGAGTTCTGTGGGTCCAAAGGGCCCTCGCTGAAACACGACTATGTGA	4711
Db	3785	ATCCTGCGCGCTATGAGTTCTGTGGGTCCAAAGGCTCTGGCTGAAACACGACTATGTGA	3844
Qy	4712	AAGTCTTGAGTATGATCAAGGTCAAGTGCAGAGTTGCGTTTTCTTCCATCCCTGC	4771
Db	3845	AAGTCTTGAGACATGGTTCAGGGTCAATGCAAGAGTTGCGCATGCGCTACCCATCCCTGC	3904
Qy	4772	GTGAAGCAGCTTTTGAGAGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCCAAGGCCA	4831
Db	3905	GTGAAGCAGCTTTGTTAGAGGAGAGAGGGAGTCTGAGCATGAGTTGCAGCCAGGCGTG	3964
Qy	4832	GTGGGAGG- ---GGACTGGCCAGTGCACCTTCAGGGCCGCGTCCAGCAGCTTCCCTC	4887
Db	3965	TGGGAAGGGCCAGGGCTGGCCAGTGCATCTAAACA--GCCCTGTGCAGCAGCTTCCCTT	4022
Qy	4888	GCCTGTGTACATGAGGCCCATCTTCTCACTC---TGAAGAGAGCGGTGAGTCTGTCA	4943
Db	4023	GCCTGTGTACATGAGGCCCATCTTCTCACTCTGTTTGAAGAAATAGTCAAGTGTCTTA	4082
Qy	4944	GTAGTAGGTTCTGTTCTATTTGGGTGACTTTGGAGATTATCTTTGTTCTCTTTTGGAAAT	5003
Db	4083	GTAGTAGGTTCTATTTTGTGTGATGACTTTGGAGATTATCTCTGTTCTCTTTTACAAT	4142
Qy	5004	GTTCAATGTTTTTTTTTAAGSGATGGTTGAATGACTTACGATCCAGCATCCAGTTTATGATG	5063
Db	4143	GTTGAATG-TTCCTTTAATGGATGGTTGAATTAACCTTACGATCCAGTTTATGATG	4201
Qy	5064	ACAGCAGTCAACAGTTCTGTGTATATAGTTTAAAGGGTAAGAGTCTTGTGTTTATTCAG	5123
Db	4202	GTAGTTAACCTATATTGCTGTTAATATAGTTTAGGAGTAAGAGTCTGTTTTTATTCAG	4261
Qy	5124	ATTGGGAAATCCATTTCTATTTTGTGNAITGGG- ---ATAATACAGCAGTGGAAATAGTAC	5180
Db	4262	ATTGGGAAATCCGTTCTATTTTGTGNAITTTGGGACATAATACAGCAGTGGAGTAAGTAT	4321
Qy	5181	TTAGAAATGTGA AAAATAGACAGTAAATAGATGAGATAAAGAACTAAAGAAATTAAGAG	5240
Db	4322	TTAGAGTGTG; --AATTCACCTGNAATAGTGTGAGAT-----AAATTAAG	4366
Qy	5241	ATAGTCAATTTCTGCCTTATFACCTCAGTCTATTTCTGTA AAAATTTTTTAAAGATATATGCAT	5300
Db	4367	ATACTTAATTCGCGCCTTATGCTCAGTCTATTTCTGTA AAAATTTTTTAAAGATATATATGCAT	4426











[illegible]



Db 3552 CTGGTCTTTGGCATGTGACGTAAAGGAAGCGGAGCCACACAGCAACCTACACCTTGTCT 3611  
QY 4397 ACCTGCTAGTCTCTCTCTATGCTGGCTGCTGGTG--ATAATCAGATCATCCCAAG 4453  
Db 3612 ACCTGCTGG--ACTCTATGCTGGCTGCTGGTTGATATATCAGATCATGCCCAAG 3669  
QY 4454 ACAGGCTTCCTGATTAATGTCTCTGCTCATGATGCAATGAGGCGGCCCATGCTCTGAG 4513  
Db 3670 ACAGGCTTCCTGATTAATGCTCTGGCATGATGCAATGAGGCGCAATGCTCTGAG 3729  
QY 4514 GAGAAATCTGGGAGGAGCTGATGCTGATGAGGCTGATGATGAGGAGGACACATGCC 4573  
Db 3730 GAGAAATCTGGGAGGAGCTGATGCTGATGAGGCTGATGATGAGGAGGACACATGTC 3789  
QY 4574 TATGGGAGGCCAGGAAGCTGCTCACCACAGATTTGGTGCAGGAAAAGTACTCGAGTA- 4633  
Db 3790 TGTGGGAGGCCAGGAAGCTGCTCACCACAGATTTGGTGCAGGAAAAGTACTCGAGTAC 3849  
QY 4533 CGGAGGTGGCGGACAGTATCCCGACGCTATGAGTTCCTGTTGGGTCCAAAGGCCCTC 4692  
Db 3850 CGGAGGTGGCGGACAGTATCCCATATGCTATGAGTTACTGTGGGTCCAAAGGCCACTC 3909  
QY 4593 GCTGAACACAGCTATGTAAGTCTTTGAGTATGATCAAGTCAAGTTCAGTTCGC 4752  
Db 3910 GCTGCT-----TGAAAGTACTGGACACGCTGTGACGGTCAATGCAAGAGTTCTC 3959  
QY 4753 TTTTCTTCCCATCCCTCGGTGAAGCAGCTTTGAGAGAGGAGGAGGAGTCTGAGCA 4812  
Db 3960 ATTTCTTACCCTCCCTCGGTGAAGCAGCTTTGAGAGAGGAGGAGGAGTCTGAGCA 4019  
QY 4813 TGAGTTCAGCCAAAGGCGAGTGGGAGGGGAGTGGGCGACGTGCGACCTTCCAGGGCCGCT 4872  
Db 4020 TGAGTTCAGCCAGGCGCACTGCGAGGGGGCTGGGCGAGTGCACCTTCCAGGGCTCCGT 4079  
QY 4873 CCAGCAGCTTCCCTGCTC--TCGTGTGACATGAGGCCCATCTTTCACATC--TGAGAGAGC 4929  
Db 4080 CCAGTATTTCCCTGCTTAAATGTGACATGAGGCCCATCTTCTCTCTTTGAAGAGAGC 4139  
QY 4930 GGTGAGTGTCTCAGTAGTAGTGTCTTCTTATTTGGGTGACTTGGAGATTTATCTTTGT 4989  
Db 4140 AGTCAACATCTTAGTAGTGGTGTCTCTATTTGATGATGATGATGATGATGATGATGAT 4199  
QY 4990 TCTCTTTTGAATGTTCAATGTTTTTTTTTAAGGATGGTGAATGAATTCAGCATC 5049  
Db 4200 TTTCTTTTGAATGTTCAATG--TTTCTTTTATGAGTGTGTTGAATGAATTCAGCATC 4258  
QY 5050 CAAGTTTATGAATGACAGCAGTACAC--AGTCTGTGTATATAGTTTAAGGGTAAGAGT 5107  
Db 4259 CAATTTTATGAATGACAGTACATAGTGTCTGTTTATATAGTTTAGGAGTAAGAGT 4318  
QY 5108 CTGTGTTTTTATTCAGATTTGGGAATCCATCTTATTTTGTGAATTTGG--ATAATAACAG 5165  
Db 4319 CTGTGTTTTTATTCAGATTTGGGAATCCATCTTATTTTGTGAATTTGGACATAGTTACAG 4378  
QY 5166 CAGTGGAAATAGTACTTA--GAAATGTGAATAAGTGAAGCAATAGATGAGATAAGAA 5224  
Db 4379 CAGTGGAAATAGTACTTA--GAAATGTGAATAAGTGAAGCAATAGATGAGATAAGAA 4432  
QY 5225 CTAAGAAATTAAGAGATAGTCAATTTCTTGGCTTATACCTTATCTGTTAAATTTT 5284  
Db 4432 -TAAAGAAATTAAGAGATAGTCAATTTCTTGGCTTATACCTTATCTGTTAAATTTT 4490  
QY 5285 TTAAGATATATGATACCTGGAATTTCTTGGCTTCTTTGAGATGTAAGAGAAATTTAA 5344  
Db 4491 TTTAAATAATGTCATACCTGGAATTTCTTGGCTTCTTTGAGATGTAAGAGAAATTTAA 4550  
QY 5345 TCTGAATAAGAAATTTCTTCTGTTTCACTGGCTTTTCTTCTTCACTGAGCATCTG 5404  
Db 4551 TCTGAATAAATCATCTTCTTCTGTTTCACTGGCTTATTTATTTCTTATGCACTGAGCATTTG 4610  
QY 5405 CTTTTTGGAGGCGCTGGGTAGTAGTGGAGATGCTAAGTGAAGCCAGACTCATACCCAC 5464  
Db 4611 CTCTGTGGAAGGCGCTGGGTAAATAGTGGAGATGCTAAGGTAAGCCAGACTCACCCCTAC 4670

QY 5465 CCATAGGTCGTAGAGTCTAGGAGCTCAGTCACTAGTAACTAGAGTGGCAAGATGCTCTCT 5524  
Db 4671 CCACAGGTAGTAAAGTCTAGGAGCAGCAGTATATATTAAGTGGAGAGATCCCTCTCT 4730  
QY 5525 AA 5526  
Db 4731 AA 4732

RESULT 10  
LOCUS 136922  
DEFINITION Sequence 7 from patent US 5612201. PAT 21-APR-1997  
ACCESSION 136922  
NID 92084882  
VERSION 136922.1 GI:2084882  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2419)  
AUTHORS De Plaen, E., Boon-Falleur, T., Lethe, B., Szikora, J., De Smet, C. and Chomez, P.  
TITLE Isolated nucleic acid molecules useful in determining expression of a tumor rejection antigen precursor  
JOURNAL Patent: US 5612201-A 7 18-MAR-1997;  
FEATURES Location/Qualifiers  
source 1..2419  
BASE COUNT 562 a 581 c 677 g 599 t  
ORIGIN

Query Match : 42.6%; Score 2419; DB 5; Length 2419;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3256 GGATCCAGGCCCTGCCAGGAAATATAAGGCCCTCGTGAGAACAGAGGGGTGATCC 3315  
Db 1 GGATCCAGGCCCTGCCAGGAAATATAAGGCCCTCGTGAGAACAGAGGGGTGATCC 60  
QY 3316 ACTGCATGAGAGTGGGATGTCACAGAGTCCAGCCACCTCTCTGTAGCACTGAGAAGC 3375  
Db 61 ACTGCATGAGAGTGGGATGTCACAGAGTCCAGCCACCTCTCTGTAGCACTGAGAAGC 120  
QY 3376 CAGGGCTGCTTGGCGTCTGCACCTTGAGGCCCTGGATTCTCTCTGTAGAGTCCA 3435  
Db 121 CAGGGCTGCTTGGCGTCTGCACCTTGAGGCCCTGGATTCTCTCTGTAGAGTCCA 180  
QY 3436 GGAACAGGAGTGGAGCCCTTGGTCTGAGACAGTATCTCAGGTCACAGAGCAGAGGATG 3495  
Db 181 GGAACAGGAGTGGAGCCCTTGGTCTGAGACAGTATCTCAGGTCACAGAGCAGAGGATG 240  
QY 3496 CACAGGCTGTGCCAGCAGTGAATGTTTCCCTGAATGCACACCAAGGCCCTGACCTGCCA 3555  
Db 241 CACAGGCTGTGCCAGCAGTGAATGTTTCCCTGAATGCACACCAAGGCCCTGACCTGCCA 300  
QY 3556 CAGACACATAGAGTCCACAGAGTCTGGCCCTCACCTTCTCTCTCTCTCTCTCTCTCTCT 3615  
Db 301 CAGACACATAGAGTCCACAGAGTCTGGCCCTCACCTTCTCTCTCTCTCTCTCTCTCTCT 360  
QY 3616 CGACCTCTGCTGGCGGCTGTACCTCAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3675  
Db 361 CGACCTCTGCTGGCGGCTGTACCTCAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
QY 3676 GGGACAGGCCAACCCAGAGGAGGATTCCTTGGAGGCCACAGAGGAGCACCAGAGGAA 3735  
Db 421 GGGACAGGCCAACCCAGAGGAGGATTCCTTGGAGGCCACAGAGGAGCACCAGAGGAA 480  
QY 3736 GATCTGTAAGTAGGCTTTTGTAGAGTCTCAAGGTTTCAAGTCTCAGCTGAGGCTCTCA 3795  
Db 481 GATCTGTAAGTAGGCTTTTGTAGAGTCTCAAGGTTTCAAGTCTCAGCTGAGGCTCTCA 540



```
FEATURES          Location/Qualifiers
source            1. .2420
BASE COUNT       562 a 582 c 677 g 599 t
ORIGIN

Query Match      42.4% Score 2408; DB 5; Length 2420;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 3256 GGATCCAGGCCCTGCCAGGAAAATATAAGGGCCCTGGTGAGACAGAGGGGTCTATCC 3315
Db 1 GGATCCAGGCCCTGCCAGGAAAATATAAGGGCCCTGGTGAGAACAGAGGGGTCTATCC 60

Qy 3316 ACTGCATGAGAGTGGGATGTCACAGAGTCCAGCCACCCCTCTGGTAGCACTGAGAGC 3375
Db 61 ACTGCATGAGAGTGGGATGTCACAGAGTCCAGCCACCCCTCTGGTAGCACTGAGAGC 120

Qy 3376 CAGGGCTGTGCTTGGGCTGCACCCCTGAGGGCCCGTGGATTCTCTTCTGGAGCTCCA 3435
Db 121 CAGGGCTGTGCTTGGGCTGCACCCCTGAGGGCCCGTGGATTCTCTTCTGGAGCTCCA 180

Qy 3436 GGAACACAGGAGTGGGCTTGGTCTGAGACAGTATCTCAGGTCAACAGCAGAGATG 3495
Db 181 GGAACACAGGAGTGGGCTTGGTCTGAGACAGTATCTCAGGTCAACAGCAGAGATG 240

Qy 3496 CACAGGGTGTCCAGCAGTGAATGTTGCCCTGAATGCACACCAAGGCCCCACCTGCCA 3555
Db 241 CACAGGGTGTCCAGCAGTGAATGTTGCCCTGAATGCACACCAAGGCCCCACCTGCCA 300

Qy 3556 CAGGACATAGGACTCCACAGAGTGTGGCTCACCTCCCTACTGTCTAGTCTCTAGAA 3615
Db 301 CAGGACATAGGACTCCACAGAGTGTGGCTCACCTCCCTACTGTCTAGTCTCTAGAA 360

Qy 3616 CGACTCTGCTGGCGGTGTACCTCAGTACCTCTCAGTCTCTCTCTCTCTCTCTCT 3675
Db 361 CGACTCTGCTGGCGGTGTACCTCAGTACCTCTCAGTCTCTCTCTCTCTCTCTCT 420

Qy 3676 GGGACAGGCCAACCCAGAGGACAGGATTCCTGGAGGCCACAGAGGACACCAAGGAA 3735
Db 421 GGGACAGGCCAACCCAGAGGACAGGATTCCTGGAGGCCACAGAGGACACCAAGGAA 480

Qy 3736 GATCTGTAAAGTAGGCTTTGTTAGAGTCTCAAGGTTTCAGTTCTCAGTGAGGCTCTCA 3795
Db 481 GATCTGTAAAGTAGGCTTTGTTAGAGTCTCAAGGTTTCAGTTCTCAGTGAGGCTCTCA 540

Qy 3796 CACACTCCCTCTCTCCAGGCTGTGGGTCTTCAATGCCAGCTCTCTGCCACACTCT 3855
Db 541 CACACTCCCTCTCTCCAGGCTGTGGGTCTTCAATGCCAGCTCTCTGCCACACTCT 600

Qy 3856 GCCTGTGCTGCTGACGAGTCAATCTCTCTGAGCAGAGGAGTCTGCACTGCAAGCC 3915
Db 601 GCCTGTGCTGCTGACGAGTCAATCTCTCTGAGCAGAGGAGTCTGCACTGCAAGCC 660

Qy 3916 TGAGGAAGCCCTTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCAC 3975
Db 661 TGAGGAAGCCCTTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCAC 720

Qy 3976 CTCCTCTCTCTCTCTCTGCTGCTGGGACCCCTGGAGAGGTGCCACTGCTGGGTCAAC 4035
Db 721 CTCCTCTCTCTCTCTCTGCTGCTGGGACCCCTGGAGAGGTGCCACTGCTGGGTCAAC 780

Qy 4036 AGATCTCCCCAGAGTCTCTCAGGAGGCTCTGGCTTTCCCACTACCATCTCACTCG 4095
Db 781 AGATCTCCCCAGAGTCTCTCAGGAGGCTCTGGCTTTCCCACTACCATCTCACTCG 840

Qy 4096 ACAGAGCAACCCAGTGGGTTCCAGCAGCCGTGAAGAGAGGGGCCAAGCACTCTTG 4155
Db 841 ACAGAGCAACCCAGTGGGTTCCAGCAGCCGTGAAGAGAGGGGCCAAGCACTCTTG 900

Qy 4156 TATCCTGGAGTCTCTGTCAGCAGTAACTCACTAAGAGGTGCTGATTTGGTGGTTT 4215
Db 1561 TATCCTGGAGTCTCTGTCAGCAGTAACTCACTAAGAGGTGCTGATTTGGTGGTTT 960

Qy 4216 TCTGCTCTCAAAATATCGAGCCAGGAGCCAGTCAACAAGGACGAAATGCTGGAGAGTGT 4275
Db 961 TCTGCTCTCAAAATATCGAGCCAGGAGCCAGTCAACAAGGACGAAATGCTGGAGAGTGT 1020

Qy 4276 CATCAAAAATTTACAAGCACTGTTTTCTGAGATCTTCGGCAAAAGCCCTCTGAGTCTGCA 4335
Db 1021 CATCAAAAATTTACAAGCACTGTTTTCTGAGATCTTCGGCAAAAGCCCTCTGAGTCTGCA 1080

Qy 4336 GCTGGTCTTTGGCAATTCAGCTGAAGGAGCAGACCCACCAGGCTCTTATGTCCTTGT 4395
Db 1381 GCTGGTCTTTGGCAATTCAGCTGAAGGAGCAGACCCACCAGGCTCTTATGTCCTTGT 1140

Qy 4396 CACCTGCTAGTCTCTCTCTATGATGCTCTGGGTGATTAATCAGATCATGCCCAAGAC 4455
Db 1141 CACCTGCTAGTCTCTCTCTATGATGCTCTGGGTGATTAATCAGATCATGCCCAAGAC 1200

Qy 4456 AGGCTTCTGATAATTCCTGCTCATGATTGCAATGGAGGGCCGCTCTCTCTGAGGA 4515
Db 1201 AGGCTTCTGATAATTCCTGCTCATGATTGCAATGGAGGGCCGCTCTCTCTGAGGA 1260

Qy 4516 GGAATCTGGAGGAGCTGAGTGTGATGGAGGTGATGATGGGAGGAGCAGTGCCTA 4575
Db 1261 GGAATCTGGAGGAGCTGAGTGTGATGGAGGTGATGATGGGAGGAGCAGTGCCTA 1320

Qy 4576 TGGGAGCCAGGAGGCTGCTCACCAAGATTGCTGTCAGGAAAGTACCTGGAGTA--CG 4634
Db 1321 TGGGAGCCAGGAGGCTGCTCACCAAGATTGCTGTCAGGAAAGTACCTGGAGTACCG 1380

Qy 4635 GCAGGTGCCGGACAGTATCCGACAGCTATGATGCTCTGTTGGGTCCAAGGGCCCTCGC 4694
Db 1381 GCAGGTGCCGGACAGTATCCGACAGCTATGATGCTCTGTTGGGTCCAAGGGCCCTCGC 1440

Qy 4695 TGAACACAGTATGTGAAGTCTTGTGATGATGATCAAGGTCTGATGCAAGAGTTCGCTT 4754
Db 1441 TGAACACAGTATGTGAAGTCTTGTGATGATGATCAAGGTCTGATGCAAGAGTTCGCTT 1500

Qy 4755 TTTCTTCCCATCCCTGCTGAAGCAGCTTTGAGAGAGGAGGAGGAGTCTGAGCATG 4814
Db 1501 TTTCTTCCCATCCCTGCTGAAGCAGCTTTGAGAGAGGAGGAGGAGTCTGAGCATG 1560

Qy 4815 AGTTGACGCAAGGCCAGTGGAGGGGAGTGGCCAGTGCACCTTCCAGGGCCGCGTCC 4874
Db 1561 AGTTGACGCAAGGCCAGTGGAGGGGAGTGGCCAGTGCACCTTCCAGGGCCGCGTCC 1620

Qy 4875 AGCAGCTTCCCTGCTCGTGACATGAGGCCCATTTCTACCTCTGAAGAGAGCGGTCA 4934
Db 1621 AGCAGCTTCCCTGCTCGTGACATGAGGCCCATTTCTACCTCTGAAGAGAGCGGTCA 1680

Qy 4935 GTGTTCTCAGTAGTGGTCTTCTGTTCTATTGGGTGACTTGGAGATTATCTTCTCTCT 4994
Db 1681 GTGTTCTCAGTAGTGGTCTTCTGTTCTATTGGGTGACTTGGAGATTATCTTCTCTCT 1740

Qy 4995 TTTGGAATTTTCAAAATGTTTTTTTTTAAGGATGTTGTAATGAATCTCAGCATCCAAGT 5034
Db 1741 TTTGGAATTTTCAAAATGTTTTTTTTTAAGGATGTTGTAATGAATCTCAGCATCCAAGT 1800

Qy 5055 TTATGAATGACAGCAGTCAACAGTCTCTGTGATATAGTTTAAAGGTAAGAGTCTTGTGT 5114
Db 1801 TTATGAATGACAGCAGTCAACAGTCTCTGTGATATAGTTTAAAGGTAAGAGTCTTGTGT 1860

Qy 5115 TTTATTTCAGATTGGGAATCCATTCTATTGTAATGGGTAATTAACAGCAGTGGAA 5174
Db 1861 TTTATTTCAGATTGGGAATCCATTCTATTGTAATGGGTAATTAACAGCAGTGGAA 1920

Qy 5175 AAGTACTTAGAAATGTGAATAAGAGCAGTAAATAGATGAGATAAGAACTAAAGAAAT 5234
Db 1921 AAGTACTTAGAAATGTGAATAAGAGCAGTAAATAGATGAGATAAGAACTAAAGAAAT 1980

Qy 5235 TAAGAGATGATCAATCTTCCCTTATACCTCAGTCTATCTGTAATAATTTTTAAGATAT 5294
Db 1981 TAAGAGATGATCAATCTTCCCTTATACCTCAGTCTATCTGTAATAATTTTTAAGATAT 2040
```



```

QY 5295 ATGCATACCTGGATTCCTTCCTGGCTCTTTGAGATGTAAGAGAAATTAATCTGAATAA 5354
Db 2041 ATGCATACCTGGATTCCTTCCTGGCTCTTTGAGATGTAAGAGAAATTAATCTGAATAA 2100
QY 5355 GAATTCCTCTGCTCACTGCTCTTTCTTCCTCATGCACTGAGCATCTCTTTTGGAA 5414
Db 2101 GAATTCCTCTGCTCACTGCTCTTTCTTCCTCATGCACTGAGCATCTCTTTTGGAA 2160
QY 5415 GGCCTGGGTTAGTAGGGAGATGCTAAGGTAAAGCCAGACTCATACCCACCATAGGTC 5474
Db 2161 GGCCTGGGTTAGTAGGGAGATGCTAAGGTAAAGCCAGACTCATACCCACCATAGGTC 2220
QY 5475 GTAGAGTCTAGGAGCTGACGTACGTAAATCAGAGTGTCAGAGTGTCTCTTAAGAGTAG 5534
Db 2221 GTAGAGTCTAGGAGCTGACGTACGTAAATCAGAGTGTCAGAGTGTCTCTTAAGAGTAG 2280
QY 5535 GGAAGAGTAGAGAGGGGTGAGGGTGTGGGCTCCGGGTGAGAGTGTGAGTGTCAATG 5594
Db 2281 GGAAGAGTAGAGAGGGGTGAGGGTGTGGGCTCCGGGTGAGAGTGTGAGTGTCAATG 2340
QY 5595 CCTGAGCTGGGCACTTTGGGCTTTGGGAACTGCAAGTTCCTCTCTGGGGAGCTGATTG 5654
Db 2341 CCTGAGCTGGGCACTTTGGGCTTTGGGAACTGCAAGTTCCTCTCTGGGGAGCTGATTG 2400
QY 5655 TAATGATCTTGGGTGGATCC 5674
Db 2401 TAATGATCTTGGGTGGATCC 2420

RESULT 12
LOCUS HUMMAG1A 2420 bp DNA PRI 15-NOV-1993
DEFINITION Human antigen (MAGE-1) gene, complete cds.
ACCESSION M77481
NID 9416114
VERSION M77481.1 GI:416114
KEYWORDS antigen.
SOURCE Homo sapiens (individual isolate patient M22) melanoma abdominal
metastasis of melanoma DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 785 to 1286)
AUTHORS van der Bruggen,P., Traversari,C., Chomez,P., Lurquin,C., De
Plaen,E., Van den Eynde,B., Knuth,A. and Boon,T.
TITLE A gene encoding an antigen recognized by cytolytic T lymphocytes on
a human melanoma
JOURNAL Science 254, 1643-1647 (1991)
MEDLINE 92086861
REFERENCE 2 (bases 1 to 2420)
AUTHORS van der Bruggen P.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-1992) Pierre van der Bruggen, Ludwig Institute
for Cancer Research, Brussels Branch, Avenue Hippocrate, 74, UCL
7459, Brussels, B-1200, Belgium
COMMENT On Nov 15, 1993 this sequence version replaced gi:187294.
FEATURES
Source
1..2420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="x"
/sex="female"
/cell_line="M22-MEL.43"
1..412
/partial
/gene="MAGE-1"
/number=1
413..485
/gene="MAGE-1"
/number=2
join(413..485,561..2111)
/partial
Intron
exon
mRNA

```

```

/ gene="MAGE-1"
486..560
/ gene="MAGE-1"
/ number=2
561..2111
/ gene="MAGE-1"
/ number=3
626..1555
/ gene="MAGE-1"
626..1555
/ gene="MAGE-1"
/codon_start=1
/product="MAGE-1"
/protein_id="AA03229.1"
/db_xref="PI:416115"
/db_xref="GI:416115"
/translation="MSLEORSLHCKPEEALQAQOEALGLVCVQAATSSSSPLVLGTLTLE
SVPTAGSTDPPOSPOGASAPFTTINFRQPSSESGSREEGSPSCILESIFRAVI
TKKVADLVGLLKYRAREPVTKAEMLESVIRKYNKHFPEIFGKASESLQLVFGIDVK
EADPTCHSYLVLTCLGSLDGLDNQIMPKTGFLLIIVLMIAAMEGSHAPEEIEWEL
SYMEVYDGRHSAYGEPRKLLTODLVOEKYLEYRQVPDSDPARYEFLWGPRALETSY
VKVLEVYIKVSARVRFPPFSLREALREEEGV"
polyA_signal 2095..2100
/ gene="MAGE-1"
2112..2420
/partial
/ gene="MAGE-1"
/ number=3
BASE COUNT 562 a 582 c 677 g 599 t
ORIGIN
Query Match 42.4%; Score 2408; DB 9; Length 2420;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 3256 GGATCAGGCCCTGCCAGGAAAAATATAAGGGCCCTCGGTGAGAACAGAGGGGTCTATCC 3315
Db 1 GGATCAGGCCCTGCCAGGAAAAATATAAGGGCCCTCGGTGAGAACAGAGGGGTCTATCC 60
QY 3316 ACTGCATGAGAGTGGGATGTACAGAGTCCAGCCACCTCTCTGTAGCACTGAGAGC 3375
Db 61 ACTGCATGAGAGTGGGATGTACAGAGTCCAGCCACCTCTCTGTAGCACTGAGAGC 120
QY 3376 CAGGGCTGTCTTGGCTGTGACCTGAGGGCCGCTGGATTCTCTCTGTGAGCTCCA 3435
Db 121 CAGGGCTGTCTTGGCTGTGACCTGAGGGCCGCTGGATTCTCTCTGTGAGCTCCA 180
QY 3436 GGAACAGGCAGTGAAGCCCTTGGTCTGAGACAGTATCCTCAGGTACAGAGCAGAGATG 3495
Db 181 GGAACAGGCAGTGAAGCCCTTGGTCTGAGACAGTATCCTCAGGTACAGAGCAGAGATG 240
QY 3496 CACAGGTGTGCCAGCAGTGAATGTTTGCCTGAATGCACACCAAGGCCCCACCTGCCA 3555
Db 241 CACAGGTGTGCCAGCAGTGAATGTTTGCCTGAATGCACACCAAGGCCCCACCTGCCA 300
QY 3556 CAGGACACATAGGACTCCACAGAGTCTGGCTCACCTCCCTACTGTCTGCTGCTAGTAAT 3615
Db 301 CAGGACACATAGGACTCCACAGAGTCTGGCTCACCTCCCTACTGTCTGCTGCTAGTAAT 360
QY 3616 CGACCTCTGCTGGCCGGCTGTACCTGAGTACCTCTCACTTCTCTCTCAGGTTTTCAG 3675
Db 361 CGACCTCTGCTGGCCGGCTGTACCTGAGTACCTCTCACTTCTCTCTCAGGTTTTCAG 420
QY 3676 GGGACAGGCCAACCCAGAGACAGGATTCCCTGGAGCCACAGAGGACACCAAGAGAA 3735
Db 421 GGGACAGGCCAACCCAGAGACAGGATTCCCTGGAGCCACAGAGGACACCAAGAGAA 480
QY 3736 GATCTGTAAGTAGGCTTTTGTAGAGTCTCCAAGGTTCAAGTTTCTCAGCTGAGGCCCTCA 3795
Db 481 GATCTGTAAGTAGGCTTTTGTAGAGTCTCCAAGGTTCAAGTTTCTCAGCTGAGGCCCTCA 540
QY 3796 CACACTCCCTCTCTCCCCAGGCGCTGGGGTCTTCAATTGCCCAGCTCTCTGCCACACTCCT 3855

```

Db	541	CACACTCCCTCTCTCCCCAGGCGCTGGGGTCTTCATTTGCCAGTCTCTGCCACACACTCCCT	600
Qy	3856	GCCTGTGCCCCCTCAGCAGAGTTCATCATGTCTCTTGACGACAGGAGTCTGCACTCGAAGACC	3915
Db	501	GCCTGTGCCCCCTCAGCAGAGTTCATCATGTCTCTTGACGACAGGAGTCTGCACTCGAAGACC	660
Qy	3916	TGAGGAAGCCCTTTGAGGCCCAACAAGAGCCCTTGCCCTGGTGTGTGTGTCAGGCTGCCAC	3975
Db	661	TGAGGAAGCCCTTTGAGGCCCAACAAGAGCCCTTGCCCTGGTGTGTGTGTCAGGCTGCCAC	720
Qy	3976	CTCCTCTCTCTCTCTCTGGTCTTGGGCACCTCGAGAGAGTGGCCACTGCTGGGTCAAC	4035
Db	721	CTCCTCTCTCTCTCTCTGGTCTTGGGCACCTCGAGAGAGTGGCCACTGCTGGGTCAAC	780
Qy	4036	AGATCTCTCCACAGTCTCTCAGGAGCCCTCCGCCCTTCCCACTACCATCAACTTCACTCG	4095
Db	781	AGATCTCTCCACAGTCTCTCAGGAGCCCTCCGCCCTTCCCACTACCATCAACTTCACTCG	840
Qy	4096	ACAGAGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGAGGGGCCAAGCACTCTTG	4155
Db	841	ACAGAGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGAGGGGCCAAGCACTCTTG	900
Qy	4156	TATCCTGGAGTCTTGTTCGGAGCAGTAACTACTAAGAAGTGGCTGATTTGGTTGGTTT	4215
Db	901	TATCCTGGAGTCTTGTTCGGAGCAGTAACTACTAAGAAGTGGCTGATTTGGTTGGTTT	960
Qy	4216	TCGTCTCTCAAAATTCAGCAGGAGGCGAGTCAAAAGCAGAAATTCGTGGAGAGTGT	4275
Db	961	TCGTCTCTCAAAATTCAGCAGGAGGCGAGTCAAAAGCAGAAATTCGTGGAGAGTGT	1020
Qy	4276	CATCAAAATTCAGCAGTGTTTCTTGAGATCTTCGGCAAGCCCTCGAGTCTTGCA	4335
Db	1021	CATCAAAATTCAGCAGTGTTTCTTGAGATCTTCGGCAAGCCCTCGAGTCTTGCA	1080
Qy	4336	CTGGTCTTTGGATTCAGCTGAAGAAAGCAGACCCACCGGCCACTCTCTATGTCCTGT	4395
Db	1081	GCTGGTCTTTGGATTCAGCTGAAGAAAGCAGACCCACCGGCCACTCTCTATGTCCTGT	1140
Qy	4396	CACCTGCTAGGTCCTCTATGATGGCTGCTGGTGATTAATCAGATCATGCCCAAGAC	4455
Db	1141	CACCTGCTAGGTCCTCTATGATGGCTGCTGGTGATTAATCAGATCATGCCCAAGAC	1200
Qy	4456	AGGCTTCTGATAAATGTCTTGTCATGATTCGAATGGAGGCGGCCACTGCTCTGAGGA	4515
Db	1201	AGGCTTCTGATAAATGTCTTGTCATGATTCGAATGGAGGCGGCCACTGCTCTGAGGA	1260
Qy	4516	GGAAATCTGGAGAGCTGAGTGTGATGAGGTGTATGATGGGAGGAGACAGTGCCTA	4575
Db	1261	GGAAATCTGGAGAGCTGAGTGTGATGAGGTGTATGATGGGAGGAGACAGTGCCTA	1320
Qy	4576	TGGGAGCCACGAAAGCTCTCACCACAATTTGTCAGGAAAGTACCTGGAGTA-CG	4634
Db	1321	TGGGAGCCACGAAAGCTCTCACCACAATTTGTCAGGAAAGTACCTGGAGTACCG	1380
Qy	4635	GCAGGTGCCGACAGTCATCCCGACGCTATGAGTTCCTGTGGGTCCCAAGGGCCCTCGC	4694
Db	1381	GCAGGTGCCGACAGTCATCCCGACGCTATGAGTTCCTGTGGGTCCCAAGGGCCCTCGC	1440
Qy	4695	TGAACACAGCTATGTGAAAGTCTTGTAGTATGTATCAAGGTCTAGTGCAGAGTTCGCTT	4754
Db	1441	TGAACACAGCTATGTGAAAGTCTTGTAGTATGTATCAAGGTCTAGTGCAGAGTTCGCTT	1500
Qy	4755	TTTCTTCCCATCCCTGGGTGAGCAGCTTTGAGCAGGAGGAAGAGGAGTCTCAGCATG	4814
Db	1501	TTTCTTCCCATCCCTGGGTGAGCAGCTTTGAGCAGGAGGAAGAGGAGTCTCAGCATG	1560
Qy	4815	AGTTGACCAAGCCAGTGGGAGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCTCC	4874
Db	1561	AGTTGACCAAGCCAGTGGGAGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCTCC	1620
Qy	4875	AGGAGCTTCCCTCTGCCTGTGTGACATGAGGCCCATCTTCTCACTCTGAAAGAGCGGTCA	4934

Db	1621	AGCAGCTCCCTCGCTCGTGTCACATGAGGCCCAATCTCTCACTCTGAAGACGCGTCA	1680
Qy	4935	GTGTCTCAGTAGTAGGTTTCTGTCTATTGGGTGACATTGGAGATTATCTTTTCTCTCT	4994
Db	1681	GTGTTCTCAGTAGTAGGTTTCTGTCTATTGGTGACATTGGAGATTATCTTTGTTCTCT	1740
Qy	4995	TTTGGAAATGTTCCAAATGTTTTTTTTTAAGGATGGTTGAATGAACATTACAGATCCAAGT	5054
Db	1741	TTTGGAAATGTTCCAAATGTTTTTTTTTAAGGATGGTTGAATGAACATTACAGATCCAAGT	1800
Qy	5055	TTATGAATGACACGACGTACACAGTTCTCTGTGATATAGTTTAAAGGTAAGAGTCTCTGTG	5114
Db	1801	TTATGAATGACACGACGTACACAGTTCTCTGTGATATAGTTTAAAGGTAAGAGTCTCTGTG	1860
Qy	5115	TTTATTCCAGATTGGGAAATCCATCTTATTTTGTGAATGGGATATATACACAGCAGTGAAT	5174
Db	1861	TTTATTCCAGATTGGGAAATCCATCTTATTTTGTGAATGGGATATATACACAGCAGTGAAT	1920
Qy	5175	AAGTACTTAGAAATGTGAAAAATGACGAGTAAATAGATGAGATTAAGAACTAAAGAAT	5234
Db	1921	AAGTACTTAGAAATGTGAAAAATGACGAGTAAATAGATGAGATTAAGAACTAAAGAAT	1980
Qy	5235	TAAGAGATAGTCAATTCTTGCCCTTATACCTCAGTCTATTCTGTAAAAATTTTTTAAAGATAT	5294
Db	1981	TAAGAGATAGTCAATTCTTGCCCTTATACCTCAGTCTATTCTGTAAAAATTTTTTAAAGATAT	2040
Qy	5295	ATGCATACCTGGATTTCTTGGCTTCTTTTGAGAATGTAAAGAGAAATTAATCTCAATAAA	5354
Db	2041	ATGCATACCTGGATTTCTTGGCTTCTTTTGAGAATGTAAAGAGAAATTAATCTCAATAAA	2100
Qy	5355	GAATCTTCTGTTTCACTGGCTCTTTCTTCTCCATGCACGTAGCATCTGCTTTTGGAA	5414
Db	2101	GAATCTTCTGTTTCACTGGCTCTTTCTTCTCCATGCACGTAGCATCTGCTTTTGGAA	2160
Qy	5415	GGCCCTGGGTTAGTAGTGAGATGCTTAAGGTAAGCCAGACTCATACCCACCCTATAGGTC	5474
Db	2161	GGCCCTGGGTTAGTAGTGAGATGCTTAAGGTAAGCCAGACTCATACCCACCCTATAGGTC	2220
Qy	5475	GTAGAGTCTAGGAGCTGCAGTCACGTAAATCGAGGTGGCAAGATGTCTCTAAAGATGTAG	5534
Db	2221	GTAGAGTCTAGGAGCTGCAGTCACGTAAATCGAGGTGGCAAGATGTCTCTAAAGATGTAG	2280
Qy	5535	GGAAAGTGAAGAGGGGTGAGGGTGTGGGGCTCGGGGTGAGAGTGGTGAGTGTCAATG	5594
Db	2281	GGAAAGTGAAGAGGGGTGAGGGTGTGGGGCTCGGGGTGAGAGTGGTGAGTGTCAATG	2340
Qy	5595	CCCTGAGCTGGGCAATTTTGGGCTTTGGGAAACTGCAGTTCCTTCTGGGGAGCTGATTG	5654
Db	2341	CCCTGAGCTGGGCAATTTTGGGCTTTGGGAAACTGCAGTTCCTTCTGGGGAGCTGATTG	2400
Qy	5655	TAATGATCTTTGGGTGGATCC	5674
Db	2401	TAATGATCTTTGGGTGGATCC	2420

RESULT	13
HSAF002994	
LOCUS	DNA
DEFINITION	Homo sapiens cosmids Qc4G10, Qc3C7, Qc1B11 and LK1023, complete sequence.
ACCESSION	AF002994
NID	G2121297
VERSION	AF002994.1
KEYWORDS	GI:2121297
SOURCE	HTG: human
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 54155) Gloeckner, G., Rosenthal, A., Schattevoy, R., Nordsiek, G., Kioschis, P., Poustka, A. and Drescher, B. Unpublished (1997) 2 (bases 1 to 54155)
REFERENCE	
AUTHORS	
JOURNAL	
REFERENCE	



AUTHORS Schattevoy, R., Rosenthal, A., Drescher, B. and Schattevoy, R.  
TITLE Direct Submission  
JOURNAL Submitted (07-MAY-1997) Genome Analysis, Institut für Molekular  
Biotechnologie, Beutenbergstrasse 11, Jena 07745, Germany  
FEATURES  
Location/Qualifiers  
source  
1. 54155  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1. 38360  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="cosmid Qc4G10"  
source  
1154..37199  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="cosmid Qc3C7"  
repeat\_region  
1833..2115  
/rpt\_family="L1ME3A"  
/evidence-not\_experimental  
repeat\_region  
2176..2390  
/rpt\_family="L1ME3A"  
/evidence-not\_experimental  
repeat\_region  
2404..2558  
/rpt\_family="L1MB7"  
/evidence-not\_experimental  
repeat\_region  
2573..2639  
/rpt\_family="L1MB7"  
/evidence-not\_experimental  
repeat\_region  
2639..2682  
/rpt\_family="L1MB7"  
/evidence-not\_experimental  
repeat\_region  
2699..2742  
/rpt\_family="L1MB7"  
/evidence-not\_experimental  
repeat\_region  
5127..5236  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
5336..5434  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
5435..5485  
/rpt\_family="L1MB3"  
/evidence-not\_experimental  
repeat\_region  
5527..5610  
/rpt\_family="L1MA9"  
/evidence-not\_experimental  
repeat\_region  
5611..5647  
/rpt\_family="L1MA9"  
/evidence-not\_experimental  
repeat\_region  
5684..5957  
/rpt\_family="L1PA2"  
/evidence-not\_experimental  
repeat\_region  
5983..6053  
/rpt\_family="L1MA2"  
/evidence-not\_experimental  
repeat\_region  
6060..6141  
/rpt\_family="L1MA5"  
/evidence-not\_experimental  
repeat\_region  
6145..6221  
/rpt\_family="L1PA2"  
/evidence-not\_experimental  
repeat\_region  
6236..6362  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
6363..6592  
/rpt\_family="L1MA9"  
/evidence-not\_experimental  
repeat\_region  
complement(7245..7281)  
/rpt\_family="ALU"  
/evidence-not\_experimental  
complement(7282..7359)  
/rpt\_family="ALU"  
/evidence-not\_experimental

repeat\_region  
complement(7396..7455)  
/rpt\_family="ALU"  
/evidence-not\_experimental  
repeat\_region  
complement(7456..7516)  
/rpt\_family="ALU"  
/evidence-not\_experimental  
repeat\_region  
7609..7691  
/rpt\_family="L1ME3A"  
/evidence-not\_experimental  
repeat\_region  
7840..8009  
/rpt\_family="L1ME2"  
/evidence-not\_experimental  
repeat\_region  
complement(8156..8198)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(8202..8263)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(8264..8306)  
/rpt\_family="L1PA2"  
/evidence-not\_experimental  
repeat\_region  
complement(8307..8649)  
/rpt\_family="L1PA2"  
/evidence-not\_experimental  
repeat\_region  
complement(8650..9050)  
/rpt\_family="L1PA2"  
/evidence-not\_experimental  
repeat\_region  
complement(9050..9301)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(9302..9342)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(9343..9402)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(9402..9642)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(9644..9762)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(9763..10000)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(10002..10108)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(10110..10462)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(10502..11202)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(11226..11302)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(11303..11365)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(11366..11458)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(11459..11905)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(11905..12076)  
/rpt\_family="L1"  
/evidence-not\_experimental  
repeat\_region  
complement(12081..12222)  
/rpt\_family="L1"  
/evidence-not\_experimental  
misc\_feature  
12559..14659  
/evidence-not\_experimental

```
/note="GC score = 62.40 (2101bp)"
/evidence-not_experimental
13070. .48188
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="cosmid QcIB11"
13190. .13315
/note="homology = 67.50%, score = 38, counts = 21"
/rpt_type=tandem
/rpt_unit=ccccc
/evidence-not_experimental
14966. .15077
/note="BLASTN (EMBL), 1441. .1552 of emb|U18920|HSMAGE2X
Human MAGE-2 gene exons 1-4, complete cds."
/evidence-not_experimental
15854. .54155
/organism="Homo sapiens"
/db_xref="taxon:9606"
15913. .16045
/clone="cosmid LK1023"
15913. .16045
/note="BLASTN (EMBL), 1132. .1264 of emb|U10685|HS10685
Human MAGE-10 antigen (MAGE10) gene, complete cds."
/evidence-not_experimental
16303. .16415
/note="BLASTN (EMBL), 897. .1009 of emb|U10694|HS10694
Human MAGE-9 antigen (MAGE9) gene, complete cds."
/evidence-not_experimental
16628. .16837
/note="BLASTN (EMBL), 508. .717 of emb|U77481|HSMAG1A Human
antigen (MAGE-1) gene, complete cds."
/evidence-not_experimental
17057. .17199
/note="TBLASTX2 (Human exons), 2387. .2529 of
emb|U10685|HS10685 Human MAGE-10 antigen (MAGE10) gene,
complete cds."
/evidence-not_experimental
17342. .17484
/note="TBLASTX2 (Human exons), 2814. .2956 of

Query Match
  40.7%; Score 2307.6; DB 10; Length 54155;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 3479; Conservative 0; Mismatches 874; Indels 207; Gaps 44;

QY 1263 CCTCAGCCCGCAGCAGCCCAACCCCTTCTGCCACCTCACCCCTCACTGCCCCCAACCC 1322
   || || || || || || || || || || || || || || || || || || || || ||
Db 48087 CCCCAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 48146
   || || || || || || || || || || || || || || || || || || || || ||
QY 1323 CACCTCATCTCTCATGTGCCCA--CTCCCATCCGCTCCGCCCATCTTGGCAGAAATCC 1380
   || || || || || || || || || || || || || || || || || || || || ||
Db 48147 CAACACATCTTATGTCTTACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 48206
   || || || || || || || || || || || || || || || || || || || || ||
QY 1381 GGT-TGCCCTCTCTCAACCCAGGAGCCCTGGTAGCCCGATGTGAACCACTGAC 1439
   || || || || || || || || || || || || || || || || || || || || ||
Db 48207 GGTTCACCCCTTGCCTGGAACCCAGGAGAGTACGG--GCCCGGATGTGACGCCACTGC 48264
   || || || || || || || || || || || || || || || || || || || || ||
QY 1440 TTGAACCTCAGATCTGAGAGAGAGCAGGTTCATTTAATGGTTCTCAGGGGGGGCTGA 1499
   || || || || || || || || || || || || || || || || || || || || ||
Db 48265 TTGGCATTGGAGTCTAGAGAGAGCAGGATTC-----TCGCCCTGAGCAACGGCTGA 48318
   || || || || || || || || || || || || || || || || || || || || ||
QY 1500 GATCCTAGGGAGGTGGTTTATGGCTCTGTGAGGAGGAAGTGTGAGTGTGAGGGAG 1559
   || || || || || || || || || || || || || || || || || || || || ||
Db 48319 CGTCGGGGAGGGAAGCAGGCGAGGCTCGGTGAGGAGGAAGTGAAGCGCGGAGGAG 48378
   || || || || || || || || || || || || || || || || || || || || ||
QY 1560 GACTGAGGAGCAGACACCCAGGTATAGTCCCAAAATGATCCAGTACCACCCCTGCT 1619
   || || || || || || || || || || || || || || || || || || || || ||
Db 48379 GACTGAGGCGGCTTACCCGAGACAGAGGCGCCCA-ATAATCCAGCGTGCCTCTGCT 48437
   || || || || || || || || || || || || || || || || || || || || ||
QY 1620 GCCAGCCCTGGACACCCCGCCAGGAGAGATGTCTAG-----CTGGACCAACCCCGC 1672
   || || || || || || || || || || || || || || || || || || || || ||
Db 48438 GCCGGGCTTGACACACCCCTCGAGGGGAGAGTCTCTAGGCTCAGTCCGACCACTCAC 48497
   || || || || || || || || || || || || || || || || || || || || ||
QY 1673 TCCGTCCCACTGCACCTTAACCCACAGGGAATCTGTAGTCTATAGC-TTATGTGACCG 1731
   || || || || || || || || || || || || || || || || || || || || ||
```



```
Db 51772 TGGGCGAGTGCACCTTCAGGGCCCCATCCATTAGCTTCCACTGCCTCGTGATATGAG 51831
Qy 4905 GCCATTCTT--CACTCTGAAGAGAGCGGTGAGTGTCTCTCAGTAGTAGTGTCTCTGTTCTTA 4962
Db 51832 GCCATCTCTCCCTTTTGAAGAGAGAGTGCAGATCTCTTACAGTGTAGTGTCTGTTCTG 51891
Qy 4963 TTGGGTGACTTGGAGATTTATCTTTTGTCTCTTTTGGAAATGTTTCAAAATGTTTTTTT 5022
Db 51892 TTGGATGACTTTGAGATTATCTTTGTTCTCTGTTGGAATGTTTCAAAATG-TTCTTTT 51950
Qy 5023 AGGATGTTGTAAGTCACTTTCAGATCAAGTTTATGAATGACAGCAGTCAAC--AGTT 5080
Db 51951 ACAAATGTTGGATGAACTTTCAGATCAAGTTTATGAATGACAGTCAACATAGT 52010
Qy 5081 CTGTGTATATAGTTTAAAGGTTAAAGTCTCTGTTGTTTATTCAGATTTGGAAATCCATCT 5140
Db 52011 CTGTTTATATAGTTTAAAGGTTAAAGTCTCTGTTTATTCAGATTTGGAAATCCATCT 52070
Qy 5141 ATTTTGTGAATG--GGATAATAACAGCAGTGGAAATAGTACTTAGAAATCT----GAAA 5194
Db 52071 ATTTTGTGAGTTGCACATAATAACAGCAGTGGAAATAGTATTTGCCCTATATTGGAACG 52130
Qy 5195 AATGAGCAGTAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5254
Db 52131 AATTAGCAGTAAATATACATATACATGAAAGC-----TCAAAGATAGTAAATCTTG 52182
Qy 5255 CTTTATACCTCAGTCTATTCTGTAATAATTTTAAAGATATATGATGATGATGATGATG 5314
Db 52183 CTTTATACCTCAGTCTATTCTGTAATAATTTTAAAGATATATGATGATGATGATGATG 52330
Qy 5315 GGCTTTTGTGAATGTAAGAGAAATTAATCTGAATAAGAAATTTCTTCCTGTTACCTGG 5374
Db 52231 TGCTCTTTGAGATGCAAAAGAAATTAATCTGAATAAGAAATTTCTTCCTGTTACCTGG 52290
Qy 5375 CTCCTTTCTTCCATGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 5434
Db 52291 CTCATTTCTTACCATGACGATGATGATGATGATGATGATGATGATGATGATGATG 52348
Qy 5435 GATGTAAGTAAAGCAGTCACTACCCACCATAGGCTGATGATGATGATGATGATGATGATG 5494
Db 52349 GATCTAGGTAAGCAGTCACTACCCACCATAGGCTGATGATGATGATGATGATGATGATG 52408
Qy 5495 TCAGTAAATCAGGTGGCAAGATGCTCTTAAAGATGATGATGATGATGATGATGATGATGATG 5554
Db 52409 TCATGTAATTAAGTGGCGAGAGTCTCTTAGGATGATGATGATGATGATGATGATGATGATG 52468
Qy 5555 AGGCTGTTGGGCTCCGGTGGAGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGG 5614
Db 52469 AGGCTGTTGGGCTCCGGTGGAGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGG 52527
Qy 5615 GGCTTTGGGAACCTGAGTCTCTTCTGGGGAGCTGATGATGATGATGATGATGATGATGATG 5674
Db 52528 GGATCTGGGAACCTGAGTCTCTTCTGGAGAGCTGATGATGATGATGATGATGATGATGATG 52587

RESULT 14
AF030261/c 37781 bp DNA PRI 30-OCT-1997
LOCUS AF030261 Homo sapiens chromosome X clone 58F6 map Xq28, complete sequence.
DEFINITION AF030261
ACCESSION AF030261
NID 92571136
VERSION AF030261.1 GI:2571136
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 37781)
AUTHORS Gloeckner, G., Rosenthal, A., Drescher, B., Schattevoy, R., Knob, A. and
Rosenthal, A.
TITLE Sequence of cosmid 58F6 from Xq28 containing two MAGE genes
JOURNAL Unpublished
```

```
REFERENCE 2 (bases 1 to 37781)
AUTHORS Gloeckner, G., Rosenthal, A., Drescher, B., Weber, J. and Schattevoy, R.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-1997) Genome Analysis, Institut for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
Location/Qualifiers
1..37781
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xq28"
/clone="58F6"
1..5888
/notes="overlap to AF002996"
complement(2..50)
/notes="GRAIL, score = 95.000%, comment = excellent"
/evidence-not_experimental
/rpt_family="L1M4"
complement(13..561)
/evidence-not_experimental
complement(563..685)
/rpt_family="L1ME1"
/evidence-not_experimental
1198..1270
/notes="GRAIL, score = 58.000%, comment = good"
/evidence-not_experimental
2065..2232
/notes="MZE, score = 59.9%"
/evidence-not_experimental
complement(2186..2554)
/rpt_family="THE1B"
/evidence-not_experimental
2851..3207
/rpt_family="L1PA4"
/evidence-not_experimental
complement(3187..3283)
/notes="MZE, score = 73%"
/evidence-not_experimental
complement(3187..3283)
/notes="GRAIL, score = 100.000%, comment = excellent
shadow"
/evidence-not_experimental
3683..3859
/notes="GRAIL, score = 81.000%, comment = excellent"
/evidence-not_experimental
3786..3859
/notes="Xpound exon prediction, score = 92% (0%)"
/evidence-not_experimental
3939..4076
/notes="Xpound exon prediction, score = 87% (0%)"
/evidence-not_experimental
complement(5711..5857)
/notes="GRAIL, score = 90.000%, comment = excellent"
5720..5949
/notes="GRAIL, score = 68.000%, comment = good shadow"
/evidence-not_experimental
5722..5858
/notes="BLASTN2 (EST exons), 1:58F6_00_Ex_7 1..137 of
gb|U52239|HSU52239 Human chromosome X EST C49g1. 7/96 P -
1.8e-15 S = 483"
/evidence-not_experimental
5722..5858
/notes="BLASTN2 (EST exons), 1:58F6_00_Ex_8 1..137 of
gb|U52239|HSU52239 Human chromosome X EST C49g1. 7/96 P -
1.8e-15 S = 483"
/evidence-not_experimental
5761..5949
/notes="MZE, score = 56.6%"
/evidence-not_experimental
6407..6458
/notes="MZE, score = 70.3%"
/evidence-not_experimental
```

```
exon      5820..5917
/note="Xpound exon prediction, score = 72% (0%)"
/evidence=not_experimental
misc_feature 6916..37781
/note="overlap to AF002994"
misc_feature 7913..8192
/note="CpG island score = 0.86, GC = 62.10%, CpGs = 24"
/note="Region: CpG island"
/evidence=not_experimental
repeat_region 7914..8216
/rpt_family="AluYa5"
/evidence=not_experimental
misc_feature 8051..8192
/note="GC score = 11.80 (142bp)"
/note="Region: GC content"
/evidence=not_experimental
exon      complement(9430..10349)
/note="GRAIL, score = 84.000%, comment = excellent"
/evidence=not_experimental
exon      complement(9436..10349)
/note="Xpound exon prediction, score = 93% (0%)"
/evidence=not_experimental
exon      complement(9457..9712)
/note="BLASTN2 (EST exons), 1:58F6_00.Ex_11 1..256 of
gb|T29724|T29724 EST92093 Homo sapiens cDNA 5' end similar
to MAGE-2 P = 8.1e-50 S = 1238"
/evidence=not_experimental
exon      9636..10109
/note="WZEF, score = 70.5%"
/evidence=not_experimental
exon      complement(9636..9712)
/note="BLASTN2 (EST exons), 1:58F6_00.Ex_12 1..77 of
gb|T29724|T29724 EST92093 Homo sapiens cDNA 5' end similar
to MAGE-2 P = 2.7e-11 S = 385"
/evidence=not_experimental
exon      9636..9849
/note="GRAIL, score = 84.000%, comment = excellent shadow"
/evidence=not_experimental
exon      complement(9637..9731)
/note="BLASTN2 (EST exons), 2:58F6_00.Ex_12 1..95 of
gb|AA378043|AA378043 EST90692 Synovial sarcoma Homo
sapiens cDNA 5' end P = 1.7e-09 S = 343"
/evidence=not_experimental
exon      10213..10370
/note="GRAIL, score = 55.000%, comment = good shadow"
/evidence=not_experimental
exon      complement(11315..11404)
/note="WZEF, score = 71.4%"
/evidence=not_experimental
misc_feature 11759..13147
/note="GC score = 43.60 (1389bp)"
/note="Region: GC content"
/evidence=not_experimental
exon      12228..12326
/note="GRAIL, score = 74.000%, comment = good shadow"
/evidence=not_experimental
exon      12616..12673
/note="GRAIL, score = 83.000%, comment = excellent shadow"
/evidence=not_experimental
misc_feature 12682..12875
/note="CpG island score = 1.00, GC = 65.50%, CpGs = 18"
/note="Region: CpG island"
/evidence=not_experimental
exon      complement(12709..12787)
/note="GRAIL, score = 52.000%, comment = good"
/evidence=not_experimental
exon      complement(13598..13656)
/note="WZEF, score = 55.1%"
/evidence=not_experimental
exon      14076..14135
/note="WZEF, score = 99.9%"
/evidence=not_experimental
```

```
exon      14076..14135
/note="GRAIL, score = 100.000%, comment = excellent"
/evidence=not_experimental
repeat_region complement(14141..15321)
/rpt_family="L1M4"
/evidence=not_experimental
repeat_region complement(15683..15875)
/rpt_family="L1MD3"
/evidence=not_experimental
repeat_region complement(15888..16252)
/rpt_family="MLT1A2"
/evidence=not_experimental
repeat_region complement(16262..17531)
/rpt_family="L1MCI"
/evidence=not_experimental
repeat_region 17541..17701
/rpt_family="MLT1A2"
/evidence=not_experimental
repeat_region 17821..18010
/rpt_family="MLT1A2"
/evidence=not_experimental
repeat_region complement(18071..18123)
/rpt_family="L1"
/evidence=not_experimental
repeat_region 18126..18361
/rpt_family="MLT2FB"
/evidence=not_experimental
exon      18278..18377
/note="GRAIL, score = 44.000%, comment = marginal"
/evidence=not_experimental
exon      complement(19213..19255)
/note="GRAIL, score = 61.000%, comment = good shadow"
/evidence=not_experimental
repeat_region complement(19228..19766)
/rpt_family="L1M4"

Query Match      40.7%; Score 2307.6; DB 11; Length 37781;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 3479; Conservative 0; Mismatches 874; Indels 207; Gaps 44;

QY 1263 CTCTCCAGCCCGCCAGCAGCCAGCCCAACCTTCTGCCACCTCAGCCTCAGCTGCCCAACCC 1322
   || || || || || || || || || || || || || || || || || || || || || ||
DB 12983 CCCCCGACCCCGCAGGATCTACAGCTCAGGATCCCGCTCCCAATCCCTACCTACAC 12924
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1323 CACCCCTCATCTCTTCATGTGCCCA--CTGCCATCGCTCCCGCCATTTGCGAGAATCC 1380
   || || || || || || || || || || || || || || || || || || || || || ||
DB 12923 CAACACCATCTTCATGCTTACCCCAACCCCACTCCAGATCCCATCGGCGAGAATCC 12864
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1381 GGT-TGGCCCTGCTCTCAACCCAGGAGCCCTGTGTAGCCCGATGTGAAACCATGAC 1439
   || || || || || || || || || || || || || || || || || || || || || ||
DB 12863 GGTTCACCCCTTCGCTGAACCCAGGAAAGTACAGG--GCCCGGATGTGACGCCATGAC 12806
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1440 TTGAACCTCACAGATCTGAGAGAAGCCAGGTTCAATTAATGGTCTGAGGGGGGCTTGA 1499
   || || || || || || || || || || || || || || || || || || || || || ||
DB 12805 TTGCGCATTTGGAGGTTCAGAGGACAGCGAGATTC-----TCGCCCTGAGCAAGCGCTGA 12752
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1500 GATPCACTAGGGGAGTGGTTTATAGCTCTGTGAGGAGCAAGGTGAGATGTGAGGGAG 1559
   || || || || || || || || || || || || || || || || || || || || || ||
DB 12751 CGTCGCGCAGGAGCAAGCAGCGCAGGCTCCGTGAGGAGCAAGGTAAAGCGCGAGGGAG 12692
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1560 GACTGAGGAGGACACACACCCAGGTAGATGCCCCCAAAATGATCCAGTACCACCCCTGCT 1619
   || || || || || || || || || || || || || || || || || || || || || ||
DB 12691 GACTGAGGAGGAGGCTTACCCCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 12633
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1620 GCCAGCCCTGGACACACCCCGGAGGAGAGATGTCTCAG-----CTGACACACCCCGG 1672
   || || || || || || || || || || || || || || || || || || || || || ||
DB 12632 GCGGGGCTTGGACACCCCTTGCAGGGGAGAACTTCTCAGGCTAGTCCGACACCTACAC 12573
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1673 TCCGCTCCCACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGC-TTATGTGACCGG 1731
```



[illegible]

Qy	4845..	TGGGCGAGTGCACCTTCCAGGCGCGGTCACAGACGCTTCCCGCTGCCTGCTGCATGAG	4904
Db	9298	TGGGCGAGTGCACCTTCCAGGCGCGCCATCCATTATCTTCCACTGCCCTGCTGATGAG	9239
Qy	4905	GCCCATCTCT - CACTCTCAAGACGCGGTCTCAGTCTCTCAGTAGTAGGTTTCTGTCTA	4962
Db	9238	GCCCATCTCTGCGCTCTTTGAAGAGACGAGTCAGCAATCTTAGCAGTAGTTTCTGTCTG	9179
Qy	4963	TTGGGTGACCTTGAGATTTATCTTTGTCTCTTTTGGAAATGTTCAAATGTTTTTTTTTA	5022
Db	9178	TTGGATGACCTTTGAGATTTATCTTTGTTTCCGTGTTGGAATGTTCCAAATG - TTCC	9120
Qy	5023	AGGGATGGTGAATGAACCTTCAGCATCCAAAGTTTATGAATGACAGCAGTCAAC -	5080
Db	9119	ACAAATGGTGTGAATGAACCTTCAGCATCCAAAGTTTATGAATGACAGTCAACATAGT	9060
Qy	5081	CTGTGTATATAGTTTAAAGGTGAAGTCTTGTGTTTTTATTCAGATTGGGAATCCATTCT	5140
Db	9059	CTGTTTATATAGTTTAAAGGTGAAGTCTGTTTTTATTCAGATTGGGAATCCATTCC	9000
Qy	5141	ATTTTGTGAATTC - GCATAATAACAGCAGTGGAAATAGTACTTTAGAAATCT -	5194
Db	8999	ATTTGTGAGTGTGCACATAATAACAGCAGTGGAAATAGTATTTGCCCTATATCTG	8940
Qy	5195	AATGAGCAGTAATAATAGATGAGATAAAGAACTTAAAGAAATTAAGAGATAGTCAATCTTG	5254
Db	8939	AATTAGCAGTAAAAATCATGATACAAGAAC - - - - - TCAAAAGATAGTTAATCTCTG	8888
Qy	5255	CCATTACCTCAGTCTATTCTGTAAATTTTAAAGATATATGATACCTGGATTTCCTT	5314
Db	8887	CCTTATACCTCAGTCTATTATGTAAAA - - - - - TTAAAAATATGTGA - - - - -	8840
Qy	5315	GGCTCTTTGAGAAATCTAAGAGAAATTAATCTGAATAAAGAAATTCCTCCTGTTCACTGG	5374
Db	8839	TGCTCTTTGAGAAATCAAAAGAAATTAATCTGAATAAATTAATCTCCTGTTCACTGG	8780
Qy	5375	CTCTTTTCTTCCATCGACTGAGCATCTGCTTTTGGAAAGCCCTGGGTTAGTAGTGA	5434
Db	8779	CTCATTTCTTTACCATCTACTCAGCATCTGCTCTGTGGAAGCCCTGG - - TAGTAGTGGG	8722
Qy	5435	GATGCTAAGTGAAGCCAGACTCATACCCACCATAGGGTCTGTAGAGCTAGGAGCTGCAG	5494
Db	8721	GATCTAAGGTAAGCCAGACTCACGCTCTACCCATAGGCTCATAGAGTCTAGGAGCTGCAG	8662
Qy	5495	TCAGTAACTCAGGTGCGCAAGATGTCCTCTAAAGATGTAGGGAAGAGTGACAGAGGGGTG	5554
Db	8661	TCATGTAATTAAGTGGCAGAGAGTCTCTAGGATGTAGTGGAAATGTAAGCAGAGGGTG	8602
Qy	5555	AGGGTGTGGGGTCTCCGGGTGAGAGTGGTGAGTGTCATGCGCCCTGAGCTGGGGCAATTTG	5614
Db	8601	AGGGTGTGGGGTTCAGAGTGAGAGTGGT - GAGTATAAATGCCCTGAGCTGGGGCAATTTG	8543
Qy	5615	GCCTTTGGGAACTGACAGTTCCTTCTGGGGGAGCTGATGTAATGATCTTGGGTGGATCC	5674
Db	8542	GGATCTGGGAACCTGCAAGTTCCTTCTGAAGGAGCTGATTTCTAATGATCCCCGGTGGTCC	8483
RESULT 15			
HSAF002996			
LOCUS			
DEFINITION Homo sapiens cosmid, complete sequence.			
PRI			
27-MAY-1997			

ACCESSION	AF002995
NID	G2121299
VERSION	AF002995.1
KEYWORDS	GI:2121299
HTG.	
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
	Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 48574)
AUTHORS	Gloeckner, G., Rosenthal, A., Drescher, B., Schattevoy, R., Poustka and Kioschis, P.
TITLE	Genomic sequence around the CPL gene



```

JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 48574)
AUTHORS Drescher, B., Rosenthal, A., Drescher, B. and Schattevoy, R.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-1997) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
source
1. .48574
/organism="Homo sapiens"
/db_xref="taxon:9606"
162. .234
/notes="GRAIL, score = 54.000%, comment = good"
/evidence-not_experimental
162. .234
/notes="M2EF, score = 98.6%"
/evidence-not_experimental
3730. .3924
/notes="GRAIL, score = 87.000%, comment = excellent"
/evidence-not_experimental
3773. .3861
/notes="M2EF, score = 92.7%"
/evidence-not_experimental
3786. .3861
/notes="Xpound exon prediction, score = 76% (0%)
/evidence-not_experimental
5591. .6777
/notes="Xpound exon prediction, score = 89% (0%)
/evidence-not_experimental
complement(6765..7014)
/notes="GRAIL, score = 46.000%, comment = marginal shadow"
/evidence-not_experimental
6982. .7100
/notes="Alignment: EST,
AA1875771|87.4|1119|AA1574811|88.1|3103|AA3615201|96.2|
1126"
/evidence-not_experimental
7010. .7100
/notes="M2EF, score = 96.8%"
/evidence-not_experimental
7101. .9009
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
7486. .7631
/notes="Alignment: EST,
|R842671|77.6|18165|R842781|77.6|18165|AA3057141|72.3|31
95"
/evidence-not_experimental
7632. .9009
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
8297. .8464
/notes="homology = 62.50%, score = 40, counts = 84"
/rpt_type=tandem
/rpt_unit=ac
/evidence-not_experimental
8734. .8773
/notes="homology = 80.00%, score = 22, counts = 20"
/rpt_type=tandem
/rpt_unit=ac
/evidence-not_experimental
8982. .9087
/notes="Alignment: EST,
|W006691|84.0|73175|R842671|88.5|166|242|R842781|89.7|166
|242|AA157481|85.4|104185|AA1875771|88.8|120199|AA30571
41|89.7|96172|AA3615201|91.0|27104"
/evidence-not_experimental
9010. .9087
/notes="M2EF, score = 51.3%"
/evidence-not_experimental
9019. .9087
/notes="GRAIL, score = 60.000%, comment = good"
/evidence-not_experimental
9088. .9166

```

```

/notes="Alignment: EST, | Intron"
/evidence-not_experimental
9167. .9294
/notes="GRAIL, score = 100.000%, comment = excellent"
/evidence-not_experimental
9167. .9229
/notes="Alignment: EST, |AA3057141|82.5|173|234"
/evidence-not_experimental
9167. .9294
/notes="M2EF, score = 97.3%"
/evidence-not_experimental
complement(9184..9392)
/notes="GRAIL, score = 45.000%, comment = marginal shadow"
/evidence-not_experimental
9217. .9294
/notes="Xpound exon prediction, score = 97% (0%)
/evidence-not_experimental
9295. .9902
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
complement(9860..10093)
/notes="GRAIL, score = 42.000%, comment = marginal shadow"
/evidence-not_experimental
9303. .10013
/notes="Xpound exon prediction, score = 97% (0%)
/evidence-not_experimental
9303. .10053
/notes="GRAIL, score = 100.000%, comment = excellent"
/evidence-not_experimental
9303. .10053
/notes="Alignment: EST,
AA3615201|92.9|233|260|AA1574811|81.1|312|422|AA1875771|8
4.1|326|476|AA0456031|82.5|1136"
/evidence-not_experimental
9346. .9988
/notes="BLASTX (GENPEPT), 1..43 of gb|M61199|HUMCS1PA_1
Human cleavage signal 1 protein mRNA, complete cds."
/evidence-not_experimental
10054. .11216
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
11216. .11282
/notes="Alignment: EST,
AA0717011|89.6|5171|AA1875771|84.4|477|539|AA0456031|86.4
137|202"
/evidence-not_experimental
11283. .11917
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
11918. .12382
/notes="GRAIL, score = 69.000%, comment = good"
/evidence-not_experimental
11918. .12273
/notes="Alignment: EST, |AA0717011|88.2|72|417"
/evidence-not_experimental
11918. .12382
/notes="M2EF, score = 99.4%"
/evidence-not_experimental
11918. .12045
/notes="BLASTX (GENPEPT), 59..186 of gb|M61199|HUMCS1PA_1
Human cleavage signal 1 protein mRNA, complete cds."
/evidence-not_experimental
12198. .12385
/notes="Xpound exon prediction, score = 95% (0%)
/evidence-not_experimental
12200. .12283
/notes="TBLASTX2 (Human exons), 554..617 of
emb|M61199|HSCS1PA Human cleavage signal 1 protein mRNA,
complete cds."
/evidence-not_experimental
12281. .12315
/notes="BLASTX (GENPEPT), 182..216 of gb|M61199|HUMCS1PA_1
Human cleavage signal 1 protein mRNA, complete cds."

```







QY 4845 TGGGCCAGTGCACCTTCCAGGCCGCCGCTCCAGCAGCTTCCCTGCTCCTGTGACATGAG 4904  
|||||  
Db 19799 TGGGCCAGTGCACCTTCCAGGCCGCCCATCCATTAGCTTCCACTGCTCCTGTGATGAG 19858  
QY 4905 GCCCATTCTT--CACTCTGAACAGAGCGGTAGTGTCTCAGTAGTAGTCTTCTTCTA 4962  
|||||  
Db 19859 GCCCATTCTGCTCTTTGAAGAGAGCAGCTCAGCATCTTAGCAGTGAGTTCTGTCTG 19918  
QY 4963 TTGGGTGACTTGGAGATTATCTTTTCTCTTTTGGAAATGTTCAAAATGTTTTTTTTTA 5022  
|||||  
Db 19919 TTGGATGACTTTGAGATTATCTTTCTCTCTTGGAAATGTTCAAAATG--TTCCTTTTA 19977  
QY 5023 AGGGATGGTTGAATGAACCTTCAAGTCCAAAGTTTATGAATGACAGCAGTCACAC--AGTT 5080  
|||||  
Db 19978 ACAAATGGTTGGATGAACCTTCAAGTCCAAAGTTTATGAATGACAGTGCACACATAGT 20037  
QY 5081 CTGTGTATATAGTTTAAAGGTAAGAGTCTGTGTTTATTTCAGATTGGGAAATCCATTCT 5140  
|||||  
Db 20038 CTGTTTATATAGTTTAGGGTAAGAGTCTGTGTTTATTTCAGATTGGGAAATCCATTCT 20097  
QY 5141 ATTTTGTGAATG--GGATAATAACAGCAGTGGAAATAAGTACTTAGAAATGT---GAAA 5194  
|||||  
Db 20098 ATTTTGTGAGTTGTACATAATAACAGCAGTGGAAATAAGTACTTAGAAATGT---GAAA 5194  
QY 5195 AATGACAGTAAATAGATGAGATAAGAACTAAAGAAATTAAGAGATAGTCAATCTTTG 5254  
|||||  
Db 20158 AATTAGCAGTAAATACATGATACAAGAAC-----TCAAAGATAGTTAAATCTTTG 20209  
QY 5255 CCTTATACCTCAGTCTATCTCTGTAATAATTTTAAAGATATATGCATACCTGGATTCCCTT 5314  
|||||  
Db 20210 CCTTATACCTCAGTCTATCTCTGTAATAATTTTAAAGATATATGCATACCTGGATTCCCTT 5314  
QY 5315 GCCTTCTTTGAGAATGTAGAGAAATTAATCTGAATAAGAAATCTTCTCTGTCTACTGG 5374  
|||||  
Db 20258 TCCTTCTTTGAGAATGCAAAAGAAATTAATCTGAATAAGAAATCTTCTCTGTCTACTGG 20317  
QY 5375 CTCCTTTCTCTCCATGCACTGAGCATCTGCTTTTGGAGGCCCTGGGTTAGTAGTGA 5434  
|||||  
Db 20318 CTCATTTCTTTACCATTCACTCAGCATCTGCTCTGTGGAAGGCCCTGG--TAGTAGTGG 20375  
QY 5435 GATGCTAAGGTAAAGCAGACTCATACCCACCCATAGGCTCGTAGAGCTAGGAGCTGCAG 5494  
|||||  
Db 20376 GATTTAAGGTAAAGCAGACTCATACCCATAGGCTCGTAGAGCTAGGAGCTGCAG 20435  
QY 5495 TCACGTAAATCGAGGTGGCAAGATGCTCTTAAAGATGTAGGAAAGAGTCAGAGAGGGGTG 5554  
|||||  
Db 20436 TCATGTAAATTAAGGTGGCAGAGTCTCTAGGATGTAGTGAATGTAAAGACAGGGGTG 20495  
QY 5555 AGGGTGTGGGGTCCGGGTGAGAGTGGTGGAGTGAATGCCCTGAGCTGGGGCATTTTG 5614  
|||||  
Db 20496 AGGGTGTGGGGTCCAGAGTGCAGAGTGGT--GAGTATAAATGCCCTGAGCTGGGCAATTTG 20554  
QY 5615 GCCTTTGGGAAACTGCAGTCTCTTCTGGGGAGCTGATTTGATATCTTGGTGGATCC 5674  
|||||  
Db 20555 GGATCTGGGAACCTGCAGTCTCTTCTGAAGAGAGCTGATTTCTAATGATGCCGGTGGGTCC 20614

Search completed: December 10, 1999, 07:37:09  
Job time: 18018 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 1999, 02:53:33 ; Search time 189.19 Seconds  
(without alignments)  
7503.514 Million cell updates/sec

Title: US-08-819-669D-8  
Perfect score: 5674  
Sequence: 1 CCCGGGCACCACTGGCATC.....TAATGATCTTGGTGGATCC 5674  
Scoring table: IDENTITY\_NUC

Searched: 311585 seqs, 125096042 residues  
Database : N\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5674	100.0	5674	1 Q72477	Tumour rejection a
2	5672.4	100.0	5674	1 Q32352	MAGE-1 nucleic aci
3	5650.8	99.6	5724	1 Q98902	Tumour rejection a
4	2419	42.6	2419	1 Q32351	Antigen E gene. Nu
5	2419	42.6	2419	1 Q72476	Tumour rejection a
6	2415.8	42.6	2419	1 T05086	M22-MEL antigen E
7	2408	42.4	2420	1 Q72472	Tumour rejection a
8	2408	42.4	2420	1 Q85435	Human melanoma ant
9	2058.8	36.3	4157	1 Q72478	Tumour rejection a
10	2057.2	36.3	4157	1 Q32353	MAGE-2 gene. Nucle
11	1762.6	31.1	4204	1 X26974	CDNA encoding MAGE
12	1678.8	29.6	2531	1 Q72482	Tumour rejection a
13	1678.8	29.6	2531	1 Q72483	Tumour rejection a
14	1677.2	29.6	2531	1 Q32357	MAGE-4 gene. Nucle
15	1675.6	29.5	2531	1 Q32358	MAGE-41 gene. Nucl
16	1541	27.2	1691	1 V69719	Tumour rejection a
17	1504.6	26.5	2305	1 Q32361	MAGE-51 genomic DN
18	1504.6	26.5	2305	1 Q72486	Tumour rejection a
19	1504.6	26.5	2305	1 T01165	MAGE-51 gene. Dete
20	1410.8	24.9	2226	1 Q32360	MAGE-5 cDNA. Nucle
21	1410.8	24.9	2226	1 Q72485	Tumour rejection a
22	988.8	17.4	1640	1 Q72480	Tumour rejection a
23	976	17.2	1640	1 Q32355	MAGE-3 cDNA. Nucle
24	900.2	15.9	1810	1 Q32364	MAGE-8 genomic DNA
25	900.2	15.9	1810	1 Q72489	Tumour rejection a
26	900.2	15.9	1810	1 T01168	MAGE-8 gene. Deter
27	875.2	15.4	1375	1 V32716	MAGE-6 cDNA. Tumou
28	834.4	14.7	1094	1 Q67865	H6/MAGE-1 expressi
29	832.8	14.7	1084	1 Q67866	H6/MAGE-1 expressi
30	830.6	14.6	1947	1 Q32363	MAGE-7 genomic DNA
31	830.6	14.6	1947	1 Q72484	Tumour rejection a
32	830.6	14.6	1947	1 T01167	MAGE-7 gene. Deter
33	815.2	14.4	1412	1 Q72490	Tumour rejection a
34	815.2	14.4	1412	1 T01169	MAGE-9 gene. Deter
35	812	14.3	1412	1 Q32365	MAGE-9 genomic DNA
36	760	13.4	1022	1 X40199	MAGE-4 encoding ge
37	750.8	13.2	1068	1 Q72484	Tumour rejection a
38	749.2	13.2	1068	1 Q32359	MAGE-4 cDNA. Nucle
39	583.4	10.3	1107	1 Q32367	MAGE-11 genomic DN
40	583.4	10.3	1107	1 Q72492	Tumour rejection a
41	583.4	10.3	1107	1 T01171	MAGE-11 gene. Dete
42	540.6	9.5	943	1 Q32356	MAGE-31 gene. Nucl
43	540.6	9.5	943	1 Q72481	Tumour rejection a

ALIGNMENTS

RESULT 1  
Q72477  
ID Q72477 standard; DNA: 5674 BP.  
AC Q72477;  
DT 22-JUN-1995 (first entry)  
DE Tumour rejection antigen MAGE-1 encoding DNA.  
KW Tumour rejection antigen; melanoma antigen-1; MAGE-1; MAGE-3;  
KW cancer; cytolytic T cells; antigen D; human leucocyte antigen;  
KW ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT cds 3881..4711  
/\*tag= a  
PN W09423031-A.  
PD 13-OCT-1994.  
PF 17-MAR-1994; U02877  
PR 26-MAR-1993; US-037230.  
PA (LUDW-) LUDWIG INST CANCER RES.  
PI Boon-falleur T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;  
DR WPI: 94-333192/41.  
PT New tumour rejection antigen precursor MAGE3 - useful in  
PT treatment and diagnosis of cancer  
PS Example 26; Page 59; 105pp; English.  
CC Q72477 is the DNA sequence which encodes melanoma antigen-1  
CC (MAGE-1). Another melanoma antigen MAGE-3 is encoded by Q72470,  
CC this is a tumour rejection antigen precursor. Melanomas  
CC characterised by the expression of MAGE-3 can be detected, or  
CC monitored, by contacting a test sample with an agent that can  
CC recognise MAGE-3. The melanoma can be treated by the administration  
CC of cytolytic T cells specific for the complex of antigen D (the  
CC mature rejection antigen derived from MAGE-3) and a human leucocyte  
CC antigen (esp. HLA-A1).  
SQ Sequence 5674 BP; 1276 A; 1644 C; 1569 G; 1185 T;

Query Match 100.0%; Score 5674; DB 1; Length 5674;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCCGGGGACCACTGGCATCCTCCCTCCCTTACCACCCCAATCCCTCCCTTTAGCCACCC 60  
DB 1 CCCGGGGACCACTGGCATCCTCCCTCCCTTACCACCCCAATCCCTCCCTTTAGCCACCC 60  
QY 61 ATCCAAACATCTTCACGCTCACGCCCAAGCCAGGCAAGTCCGGTTCACCCCTG 120  
DB 61 ATCCAAACATCTTCACGCTCACGCCCAAGCCAGGCAAGTCCGGTTCACCCCTG 120  
QY 121 CTCTCAACCACTGGAAGCCAGGTGCCAGATGTGACGCCACTGACTTGGATTAGTGG 180  
DB 121 CTCTCAACCACTGGAAGCCAGGTGCCAGATGTGACGCCACTGACTTGGATTAGTGG 180  
QY 181 TTAGAGAGAGCGAGGTTTTCGGTCTGAGGGCGCGCTTGAGATCGGTGAGGAAGCGGG 240  
DB 181 TTAGAGAGAGCGAGGTTTTCGGTCTGAGGGCGCGCTTGAGATCGGTGAGGAAGCGGG 240  
QY 241 CCCAGCTCTGTGAAGAGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300  
DB 241 CCCAGCTCTGTGAAGAGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300  
QY 301 AGATAGAGACCCCAATAATCCCTTCATGCGAGTCCCTGACCATCTGGTGGACTTC 360  
DB 301 AGATAGAGACCCCAATAATCCCTTCATGCGAGTCCCTGACCATCTGGTGGACTTC 360  
QY 361 TCAGGCTGGGCGACCCCGACAGCCCTTGTGCTTAAACCACTGGGACTCGAAGTCAGAG 420  
DB 361 TCAGGCTGGGCGACCCCGACAGCCCTTGTGCTTAAACCACTGGGACTCGAAGTCAGAG 420

QY	421	CTCCGTGTGATCAGGGAAGGGTGCTTTAGGAGAGGGCAGCGTCCAGGCTCTGCAGACAT	480
Db	421	CTCCGTGTGATCAGGGAAGGGTGCTTTAGGAGAGGGCAGCGTCCAGGCTCTGCAGACAT	480
QY	481	CATGCTCAGGATTCACAGGAGGGCTGAGGTCCTTAGACCCCACTCCCGTGACCCCAAC	540
Db	481	CATGCTCAGGATTCACAGGAGGGCTGAGGTCCTTAGACCCCACTCCCGTGACCCCAAC	540
QY	541	CCCCACTCAATGCTCACTCCCGTGACCAACCCCGCTCTTCAATTGTCTATTCCAACCCCCA	600
Db	541	CCCCACTCAATGCTCACTCCCGTGACCAACCCCGCTCTTCAATTGTCTATTCCAACCCCCA	600
QY	601	CCCCAGATCCCGACCCCATTCGCTCAACCGCTGATGCCCATCCGCCCAAGCCCATTTCCACCT	660
Db	601	CCCCAGATCCCGACCCCATTCGCTCAACCGCTGATGCCCATCCGCCCAAGCCCATTTCCACCT	660
QY	661	CACCCCAACCCCAACCCCAAGCCCACTCCGACCCCGCACCCAGGAGGATCCGGTTCCCG	720
Db	661	CACCCCAACCCCAACCCCAAGCCCACTCCGACCCCGCACCCAGGAGGATCCGGTTCCCG	720
QY	721	CCAGGAAACATCCGGGTGCCGGATGTGACGCCACTGACTTGCGCATTGTGGGCGACAGA	780
Db	721	CCAGGAAACATCCGGGTGCCGGATGTGACGCCACTGACTTGCGCATTGTGGGCGACAGA	780
QY	781	GAAGCAGGTTTCCATTCTTGAGGGAGGGGTAGATTTCGGCCGAAGAACTCGACCCAGG	840
Db	781	GAAGCAGGTTTCCATTCTTGAGGGAGGGGTAGATTTCGGCCGAAGAACTCGACCCAGG	840
QY	841	CTCTGTGAGGAGCGAGGTGAGAGCTGAGGAGGACTTGAGGACCCCGCCACTCCAATA	900
Db	841	CTCTGTGAGGAGCGAGGTGAGAGCTGAGGAGGACTTGAGGACCCCGCCACTCCAATA	900
QY	901	GAGAGCCCAAAATATTCCAGCCCGCGCTTGCTGCACAGCCCTTGCGCCACCCCGGGGAAGA	960
Db	901	GAGAGCCCAAAATATTCCAGCCCGCGCTTGCTGCACAGCCCTTGCGCCACCCCGGGGAAGA	960
QY	961	CGTCTCAGCCTGGGCTGCCCGCAGACCCCTGCTCCAAAGCCTTTGAGAGACACCAAGTTC	1020
Db	961	CGTCTCAGCCTGGGCTGCCCGCAGACCCCTGCTCCAAAGCCTTTGAGAGACACCAAGTTC	1020
QY	1021	TTCTCTCCCAAGCTCTGGAATCAGAGTTGCTGTGACCAAGGCGAGGACTGGTTAGGACAGG	1080
Db	1021	TTCTCTCCCAAGCTCTGGAATCAGAGTTGCTGTGACCAAGGCGAGGACTGGTTAGGACAGG	1080
QY	1081	GCAGGCAAGGCTCTGCCAGGCATCAAGATCAGACACCAAGAGGGAGGGCTGTGGGCC	1140
Db	1081	GCAGGCAAGGCTCTGCCAGGCATCAAGATCAGACACCAAGAGGGAGGGCTGTGGGCC	1140
QY	1141	CGAAGACTGCATCCAAATCCCACTCCACCCCATTCGCATTCGCATTCGCCACCAACC	1200
Db	1141	CGAAGACTGCATCCAAATCCCACTCCACCCCATTCGCATTCGCATTCGCCACCAACC	1200
QY	1201	CCCATCTCTCAGCTACACTCCACCCCATCCCTACTCTACTTCGGTCACTCGACCAACC	1260
Db	1201	CCCATCTCTCAGCTACACTCCACCCCATCCCTACTCTACTTCGGTCACTCGACCAACC	1260
QY	1261	ACCTTCAGCCCCAGCACCCAGCCCAACCCCTTCTGCCACCTCAACCTCACTGCCCCCAAC	1320
Db	1261	ACCTTCAGCCCCAGCACCCAGCCCAACCCCTTCTGCCACCTCAACCTCACTGCCCCCAAC	1320
QY	1321	CCACCCCTCATCTCTCATGTGCCCACTCCCATTCGCTCCCGCATTCGGGCAAGATTC	1380
Db	1321	CCACCCCTCATCTCTCATGTGCCCACTCCCATTCGCTCCCGCATTCGGGCAAGATTC	1380
QY	1381	GGTTTGCCCTTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGCATGTGAACCACTGACT	1440
Db	1381	GGTTTGCCCTTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGCATGTGAACCACTGACT	1440
QY	1441	TGAACCTCAGATCTGAGAGAAGCGAGTTCAATTAATGGTCTGAGGGGCGCTTGAG	1500
Db	1441	TGAACCTCAGATCTGAGAGAAGCGAGTTCAATTAATGGTCTGAGGGGCGCTTGAG	1500

Qy	1501	ATCCACTGAGGGAGTGGTTTTTAGGCTCTGTGTAGGAGGCAAGGTGAGATGCTGAGGGAGG	1560
Lb	1501	ATCCACTGAGGGAGTGGTTTTTAGGCTCTGTGTAGGAGGCAAGGTGAGATGCTGAGGGAGG	1560
Qy	1561	ACTGAGGAGGCACACACCCAGGTAGATGGCCCCCAAAATGATCCAGTACCACCCCTGCTG	1620
Lb	1561	ACTGAGGAGGCACACACCCAGGTAGATGGCCCCCAAAATGATCCAGTACCACCCCTGCTG	1620
Qy	1621	CCAGCCCTGACACACCCGGCCAGGACAGATGTCTCAGCTGGACACACCCCCGTCGCGTCC	1680
Lb	1621	CCAGCCCTGACACACCCGGCCAGGACAGATGTCTCAGCTGGACACACCCCCGTCGCGTCC	1680
Qy	1681	CACGTGCCACTTAACCCACAGGGCAATCTGTAGTCTCATAGTCTTATGTGACCGGGCAGGGTT	1740
Lb	1681	CACGTGCCACTTAACCCACAGGGCAATCTGTAGTCTCATAGTCTTATGTGACCGGGCAGGGTT	1740
Qy	1741	GGTCAGGAGAGGCAGGGCCAGGCATCAAGTCTCAGCATCCGCCGGCGCATTTAGGGTCAGG	1800
Lb	1741	GGTCAGGAGAGGCAGGGCCAGGCATCAAGTCTCAGCATCCGCCGGCGCATTTAGGGTCAGG	1800
Qy	1801	ACCCTGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACC	1860
Lb	1801	ACCCTGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACC	1860
Qy	1861	CCACTCACAATTCGCATACCTACCCCTACCCCAACCTCATCTTGTTCAGAACTCCCTGCTG	1920
Lb	1861	CCACTCACAATTCGCATACCTACCCCTACCCCAACCTCATCTTGTTCAGAACTCCCTGCTG	1920
Qy	1921	TCACCCACAGGAAGCCACCGGGAATGGCGCCAGGCACCTCGGATCTTGACGTCCCCATCCA	1980
Lb	1921	TCACCCACAGGAAGCCACCGGGAATGGCGCCAGGCACCTCGGATCTTGACGTCCCCATCCA	1980
Qy	1981	GGGTCTCATGAGGGAAGGGCTTGAACAGGCGCTCAGGGAGCAGAGGAGGCGCCCTAC	2040
Lb	1981	GGGTCTCATGAGGGAAGGGCTTGAACAGGCGCTCAGGGAGCAGAGGAGGCGCCCTAC	2040
Qy	2041	TGCAGATGAGGGAGGCGCTCAGAGGACCCAGCACCCCTAGGACACCGCACCCCTGTCTGAG	2100
Lb	2041	TGCAGATGAGGGAGGCGCTCAGAGGACCCAGCACCCCTAGGACACCGCACCCCTGTCTGAG	2100
Qy	2101	ACTGAGGCTGCCACTTCTGGCCTCAAGAAATCAGAACGATGGGACTCAGATTGCATGGGG	2160
Lb	2101	ACTGAGGCTGCCACTTCTGGCCTCAAGAAATCAGAACGATGGGACTCAGATTGCATGGGG	2160
Qy	2161	GTGGGACCCAGGCTCGAAGGCTTACGGCGAGGAAGAGGAGGAGGACTCAGGGAGCCTT	2220
Lb	2161	GTGGGACCCAGGCTCGAAGGCTTACGGCGAGGAAGAGGAGGAGGACTCAGGGAGCCTT	2220
Qy	2221	GGAAATCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCACAGGACACGGTGGCCACATATG	2280
Lb	2221	GGAAATCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCACAGGACACGGTGGCCACATATG	2280
Qy	2281	GCCCAATTTCTGCAATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGGGGCC	2340
Lb	2281	GCCCAATTTCTGCAATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGGGGCC	2340
Qy	2341	TCAGGTCAACAGAGGGAGGATTCAGGATCCATATGCCCCCAAGATGTGCCCTTTCATG	2400
Lb	2341	TCAGGTCAACAGAGGGAGGATTCAGGATCCATATGCCCCCAAGATGTGCCCTTTCATG	2400
Qy	2401	AGGACTGGGATATCCCGGCTCAGAAAGAAGGACTCCACACAGTCTGGCTGTCCCTT	2460
Lb	2401	AGGACTGGGATATCCCGGCTCAGAAAGAAGGACTCCACACAGTCTGGCTGTCCCTT	2460
Qy	2461	TTAGTAGCTCTAGGGGACACATCAGGATGGCGGTATGTTCCATTCCTCACTTGTACCA	2520
Lb	2461	TTAGTAGCTCTAGGGGACACATCAGGATGGCGGTATGTTCCATTCCTCACTTGTACCA	2520
Qy	2521	CAGCGCAGAAAGTTGGGGGCGCCCTCAGGAGATGGGGTCTTGGGGTAAAGGGGGATGTCT	2580
Lb	2521	CAGCGCAGAAAGTTGGGGGCGCCCTCAGGAGATGGGGTCTTGGGGTAAAGGGGGATGTCT	2580
Qy	2581	ACTCATGTCAAGGAATTTGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAT	2640

[illegible]

Db	3661	CCTTCAGGTTTTCAGGGGACAGCCGAACCCAGAGGACAGGATTCCTCGAGGCCACAGAG	3720
Qy	3721	GAGCACCAAGGAGAAGATCTGTAAGTAGGCCCTTTGTTAGAGTCTCCAAAGGTTCAAGTTCTC	3780
Db	3721	GAGCACCAAGGAGAAGATCTGTAAGTAGGCCCTTTGTTAGAGTCTCCAAAGGTTCAAGTTCTC	3780
Qy	3781	AGCTGAGGCTCTCACACACTCCCTCTCTCCCAAGGCTTGGGTCTTCATTGCCAGCT	3840
Db	3781	AGCTGAGGCTCTCACACACTCCCTCTCTCCCAAGGCTTGGGTCTTCATTGCCAGCT	3840
Qy	3841	CCTGCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAG	3900
Db	3841	CCTGCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAG	3900
Qy	3901	TCTGCACCTCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCCTGGTGTG	3960
Db	3901	TCTGCACCTCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCCTGGTGTG	3960
Qy	3961	TGTGAGGCTGCAACCTCTCTCTCTCTCTGTCCTGTCCTGGTCTGGGACCCCTGGAGAGTGCC	4020
Db	3961	TGTGAGGCTGCAACCTCTCTCTCTCTCTGTCCTGTCCTGGTCTGGGACCCCTGGAGAGTGCC	4020
Qy	4021	CACCTGCTGGGTCAACAGATCCTCCACAGAGTCCTCAGGAGCCTCCGCCTTTTCCCATAC	4080
Db	4021	CACCTGCTGGGTCAACAGATCCTCCACAGAGTCCTCAGGAGCCTCCGCCTTTTCCCATAC	4080
Qy	4081	CATCAACTTCACCTCGACAGGCAACCCAGTGAGGGTTCCAGCAGCCTGAAGAGGAGG	4140
Db	4081	CATCAACTTCACCTCGACAGGCAACCCAGTGAGGGTTCCAGCAGCCTGAAGAGGAGG	4140
Qy	4141	GCCAAGCACCTCTTGATCCTCGAGTCCTTTGTTCCGAGCAGTAATCACTAAGAAGTGTC	4200
Db	4141	GCCAAGCACCTCTTGATCCTCGAGTCCTTTGTTCCGAGCAGTAATCACTAAGAAGTGTC	4200
Qy	4201	TGATTTGGTGGTTCCTGCTCCTCAAAATATCGAGCAGGAGGCCAGTCACAAGGCCAGA	4260
Db	4201	TGATTTGGTGGTTCCTGCTCCTCAAAATATCGAGCAGGAGGCCAGTCACAAGGCCAGA	4260
Qy	4261	AATGCTGGAGAGTGTCATCAAAAATTCACAAGCACTGTTTCTCTGAGATCTTCGGCAAGC	4320
Db	4261	AATGCTGGAGAGTGTCATCAAAAATTCACAAGCACTGTTTCTCTGAGATCTTCGGCAAGC	4320
Qy	4321	CTCTGAGTCCTTGCAGCTGGTCTTTGGCAATTGACGTGAAGGAAGCAGACCCACCGGCCA	4380
Db	4321	CTCTGAGTCCTTGCAGCTGGTCTTTGGCAATTGACGTGAAGGAAGCAGACCCACCGGCCA	4380
Qy	4381	CTCCTATGTCCTTGTCACCTGCCTAGTCTCTCTATGATGGCCTGCTGGTGATATCA	4440
Db	4381	CTCCTATGTCCTTGTCACCTGCCTAGTCTCTCTATGATGGCCTGCTGGTGATATCA	4440
Qy	4441	GATCATGCCAAGCAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAANTGGAGGGCG	4500
Db	4441	GATCATGCCAAGCAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAANTGGAGGGCG	4500
Qy	4501	CCATGCTCTTGAGGAGGAATCTGGGAGGAGCTGAGTGTGATGGAGGTGATATGGGAG	4560
Db	4501	CCATGCTCTTGAGGAGGAATCTGGGAGGAGCTGAGTGTGATGGAGGTGATATGGGAG	4560
Qy	4561	GGAGCAGTGCCCTATGGGAGGCCAGGAAGCTGCTCACCNAAGATTTGGTGCAGGAAAA	4620
Db	4561	GGAGCAGTGCCCTATGGGAGGCCAGGAAGCTGCTCACCNAAGATTTGGTGCAGGAAAA	4620
Qy	4621	GTACCTGGAGTAGGAGGTCGGACAGTGATCCCGACGCTATGACTTCCTGGGGT	4680
Db	4621	GTACCTGGAGTAGGAGGTCGGACAGTGATCCCGACGCTATGACTTCCTGGGGT	4680
Qy	4681	CCAAGGGCCCTCGCTGAACCCAGCTATGTGAAGTCCCTTGAGTATGTGATCAAGGTCAGT	4740
Db	4681	CCAAGGGCCCTCGCTGAACCCAGCTATGTGAAGTCCCTTGAGTATGTGATCAAGGTCAGT	4740
Qy	4741	GCAAGAGTTCGCTTTTCTTCATCCCTGCGTGAAGCAGCTTTTGAGAGAGGAGAAAG	4800
Db	4741	GCAAGAGTTCGCTTTTCTTCATCCCTGCGTGAAGCAGCTTTTGAGAGAGGAGAAAG	4800





Db 541 CCCCACCTCCAATGCTCACTCCCGTGACCAACCCCTCTTCAATTGTATTCACACCCCA 600  
Qy 601 CCCCACATCCCCCACCACCCCTCCCTCAACCTGATGCCATCCGCCAGCCATTCACACCT 660  
Db 601 CCCCACATCCCCCACCACCCCTCCCTCAACCTGATGCCATCCGCCAGCCATTCACACCT 660  
Qy 661 CACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCC 720  
Db 661 CACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCC 720  
Qy 721 CACGGAACATCCGGGTGCCGATGTGAGCCCACTGACTTGGCCATTTGGGGCAGAGA 780  
Db 721 CACGGAACATCCGGGTGCCGATGTGAGCCCACTGACTTGGCCATTTGGGGCAGAGA 780  
Qy 781 GAAGCAGGTTTCCATTTCTGAGGAGCGGCTAGAGTTGCGCCCAAGGAACCTCACCCAGG 840  
Db 781 GAAGCAGGTTTCCATTTCTGAGGAGCGGCTAGAGTTGCGCCCAAGGAACCTCACCCAGG 840  
Qy 841 CTCGTGAGGAGCAAGGTGAGAGGCTGAGGAGGACTGAGGACCCCGCCACTCCAATA 900  
Db 841 CTCGTGAGGAGCAAGGTGAGAGGCTGAGGAGGACTGAGGACCCCGCCACTCCAATA 900  
Qy 901 GAGAGCCCAATATTTCCAGCCCGCCCTTGTCTGCCAGCCCTGGCCCAACCCCGGGAAGA 960  
Db 901 GAGAGCCCAATATTTCCAGCCCGCCCTTGTCTGCCAGCCCTGGCCCAACCCCGGGAAGA 960  
Qy 961 CGTCTCAGCCTGGGCTGCCCGCAGACCCCTGCTTCCAAAGCCCTTGAGAGACACAGGTT 1020  
Db 961 CGTCTCAGCCTGGGCTGCCCGCAGACCCCTGCTTCCAAAGCCCTTGAGAGACACAGGTT 1020  
Qy 1021 TTCTCCCCAAGCTCTGGAATCAGAGTTGCTGTGACAGGCGAGGACTGGTTAGGAGAG 1080  
Db 1021 TTCTCCCCAAGCTCTGGAATCAGAGTTGCTGTGACAGGCGAGGACTGGTTAGGAGAG 1080  
Qy 1081 GCAGGCAAGGCTCTGCCAGGATCAAGATCAGCACCACCCAGAGGAGGCTGTGGGCC 1140  
Db 1081 GCAGGCAAGGCTCTGCCAGGATCAAGATCAGCACCACCCAGAGGAGGCTGTGGGCC 1140  
Qy 1141 CCAAGACTGCACCTCAATCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCA 1200  
Db 1141 CCAAGACTGCACCTCAATCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCA 1200  
Qy 1201 CCAATCTCTCAGCTACACTCCACCCCACTCCCTACTCTCTCTCTCTCTCTCTCTCTCT 1260  
Db 1201 CCAATCTCTCAGCTACACTCCACCCCACTCCCTACTCTCTCTCTCTCTCTCTCTCTCT 1260  
Qy 1261 ACCCTCAGCCCAAGCAGCCCAACCCCTTCTGCACTCAGCTCAGCTCAGCTCAGCTCAG 1320  
Db 1261 ACCCTCAGCCCAAGCAGCCCAACCCCTTCTGCACTCAGCTCAGCTCAGCTCAGCTCAG 1320  
Qy 1321 CCAACCTCATCT 1380  
Db 1321 CCAACCTCATCT 1380  
Qy 1381 GGTTCCT 1440  
Db 1381 GGTTCCT 1440  
Qy 1441 TGAACCTCAGATCTGAGAGAGCCAGGTTCAATTTAAGGTTCTGAGGGGCGGCTTGAG 1500  
Db 1441 TGAACCTCAGATCTGAGAGAGCCAGGTTCAATTTAAGGTTCTGAGGGGCGGCTTGAG 1500  
Qy 1501 ATCCACTGAGGAGGTTTGGTCTGTGAGGAGGCAAGGTGAGTGCCTGAGGAGG 1560  
Db 1501 ATCCACTGAGGAGGTTTGGTCTGTGAGGAGGCAAGGTGAGTGCCTGAGGAGG 1560  
Qy 1561 ACTGAGGAGCAGACACCCAGGATAGTGGCCCAAAATGATCCAGTACACCCCTGCTG 1620  
Db 1561 ACTGAGGAGCAGACACCCAGGATAGTGGCCCAAAATGATCCAGTACACCCCTGCTG 1620  
Qy 1621 CAGCCCTGAGACACCCCGGCGCAGAGATGTCTCAGCTGGACACCCCGCTGCGGTC 1680  
Db 1621 CAGCCCTGAGACACCCCGGCGCAGAGATGTCTCAGCTGGACACCCCGCTGCGGTC 1680

Qy 1681 CACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTACCGGGCAGGGTT 1740  
Db 1681 CACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTACCGGGCAGGGTT 1740  
Qy 1741 GGTACGAGAGGAGGCGCCAGGCATCAAGTCCAGCATCCGCCCGGCATTTAGGGTCAG 1800  
Db 1741 GGTACGAGAGGAGGCGCCAGGCATCAAGTCCAGCATCCGCCCGGCATTTAGGGTCAG 1800  
Qy 1801 ACCCTGGAGGAACTGAGGTTCCCCACCCACACCTGTCTCTCATCTCCACCCGACC 1860  
Db 1801 ACCCTGGAGGAACTGAGGTTCCCCACCCACACCTGTCTCTCATCTCCACCCGACC 1860  
Qy 1861 CCACTCACATTTCCCATACCTACCCCTACCCCAACCTCATCTTGTGCAATCCCTGCTG 1920  
Db 1861 CCACTCACATTTCCCATACCTACCCCTACCCCAACCTCATCTTGTGCAATCCCTGCTG 1920  
Qy 1921 TCAACCCAGGAGCCACCGGAATGGCGCCAGGCATCGGATCTTGAGCTCCCATCCA 1980  
Db 1921 TCAACCCAGGAGCCACCGGAATGGCGCCAGGCATCGGATCTTGAGCTCCCATCCA 1980  
Qy 1981 GGTCTGTAGGAGGAGGCTTGACAGGSCCTCAGGGGACAGAGGAGGCGGCTTAC 2040  
Db 1981 GGTCTGTAGGAGGAGGCTTGACAGGSCCTCAGGGGACAGAGGAGGCGGCTTAC 2040  
Qy 2041 TCGCAGATGAGGAGGAGGCTCAGAGGACCCAGACACCTAGGACACCCCTGTCTGAG 2100  
Db 2041 TCGCAGATGAGGAGGAGGCTCAGAGGACCCAGACACCTAGGACACCCCTGTCTGAG 2100  
Qy 2101 ACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAACGATGGGACTCAGATTGCATGGG 2160  
Db 2101 ACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAACGATGGGACTCAGATTGCATGGG 2160  
Qy 2161 GTGGGACCCAGGCTCCAGGCTTACCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220  
Db 2161 GTGGGACCCAGGCTCCAGGCTTACCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220  
Qy 2221 GGAATCCAGATCAGTGTGAGCTCGGCCCTCAGAGGTCCAGGCGCGGTCACACATATG 2280  
Db 2221 GGAATCCAGATCAGTGTGAGCTCGGCCCTCAGAGGTCCAGGCGCGGTCACACATATG 2280  
Qy 2281 GGCATATTTCTGCACTTTTGTAGGTGACAGGACAGAGTGTGGTCTGAGAAGTGGGGCC 2340  
Db 2281 GGCATATTTCTGCACTTTTGTAGGTGACAGGACAGAGTGTGGTCTGAGAAGTGGGGCC 2340  
Qy 2341 TCAGGTCAACAGAGGAGGAGTCCAGGATCCATATGGCCCAAGATGTGCCCTTTCATG 2400  
Db 2341 TCAGGTCAACAGAGGAGGAGTCCAGGATCCATATGGCCCAAGATGTGCCCTTTCATG 2400  
Qy 2401 AGGACTGGGGATATCCCGGCTCAGAAAGAGGACTCCACACAGTCTGGCTGTCCCTT 2460  
Db 2401 AGGACTGGGGATATCCCGGCTCAGAAAGAGGACTCCACACAGTCTGGCTGTCCCTT 2460  
Qy 2461 TTAGTACTCTTAGGGGACAGATCAGGATGGCGGTATTTCCATTTCTACTTTGACCA 2520  
Db 2461 TTAGTACTCTTAGGGGACAGATCAGGATGGCGGTATTTCCATTTCTACTTTGACCA 2520  
Qy 2521 CAGGAGGAGTGGGGGCGGCTCAGGAGATGGGGTCTTGGGTAAAGGGGGATGTCT 2580  
Db 2521 CAGGAGGAGTGGGGGCGGCTCAGGAGATGGGGTCTTGGGTAAAGGGGGATGTCT 2580  
Qy 2581 ACTCATGTGAGGAAATTTGGGGTTGAGGAGCAGAGGCTGGCAGGAATTAAGATGAGT 2640  
Db 2581 ACTCATGTGAGGAAATTTGGGGTTGAGGAGCAGAGGCTGGCAGGAATTAAGATGAGT 2640  
Qy 2641 GAGACAGCAAGGCTATTGGAATCCACACCCAGAAAGAGGAGGAGGAGGAGGAGGAGG 2700  
Db 2641 GAGACAGCAAGGCTATTGGAATCCACACCCAGAAAGAGGAGGAGGAGGAGGAGGAGG 2700  
Qy 2701 TCACCCAGATGTGGCTTCTTTTCTACTCTCTTCTTCTCAGATCTGGGGCAGGTCAGG 2760  
Db 2701 TCACCCAGATGTGGCTTCTTTTCTACTCTCTTCTTCTCAGATCTGGGGCAGGTCAGG 2760



Db 4921 GAAGACAGGGTCAGTGTCTCAGTAGTAGGTTCTGTCTATTGGTGACTTGAGATT 4980  
Qy 4981 TATCTTTGTTCTCTTTTGAATTTGCAATGTTTCTTTTAAAGGATGGTTGAATGAAC 5040  
Db 4981 TATCTTTGTTCTCTTTTGAATTTGCAATGTTTCTTTTAAAGGATGGTTGAATGAAC 5040  
Qy 5041 TTCACATCCAAAGTTTATGAATGACAGCAGTCACACAGTCTGTGTATATATATTAAGG 5100  
Db 5041 TTCACATCCAAAGTTTATGAATGACAGCAGTCACACAGTCTGTGTATATATATTAAGG 5100  
Qy 5101 TAAGAGTCTGTGTTTATTTCAGATTGGAAATCAATCTATTGTAATGGGATAT 5160  
Db 5101 TAAGAGTCTGTGTTTATTTCAGATTGGAAATCAATCTATTGTAATGGGATAT 5160  
Qy 5161 AACACAGTGGAAATAGTACTTTAGAAATGTGAAATGAGCAGTAAATAGATGAGATAA 5220  
Db 5161 AACACAGTGGAAATAGTACTTTAGAAATGTGAAATGAGCAGTAAATAGATGAGATAA 5220  
Qy 5221 AGAATTAAGAAATTAAGAGATAGTCAATCTTGCTTATACCTCAGTCTATCTGTAAA 5280  
Db 5221 AGAATTAAGAAATTAAGAGATAGTCAATCTTGCTTATACCTCAGTCTATCTGTAAA 5280  
Qy 5281 ATTTTAAAGATATATGATACCTGATTTCTCTTCTGCTTCTTTGAGATGTAAGAGAAAT 5340  
Db 5281 ATTTTAAAGATATATGATACCTGATTTCTCTTCTGCTTCTTTGAGATGTAAGAGAAAT 5340  
Qy 5341 TAAATCTGAATAAAGAAATCTTCTCTTCTCAGTGGCTCTTTCTCTCCATGCACTGAGCA 5400  
Db 5341 TAAATCTGAATAAAGAAATCTTCTCTTCTCAGTGGCTCTTTCTCTCCATGCACTGAGCA 5400  
Qy 5401 TCTGCTTTTGAAGGCCCTGGGTTAGTAGTGAGATGCTAAGGTAAGCCAGCTCATAC 5460  
Db 5401 TCTGCTTTTGAAGGCCCTGGGTTAGTAGTGAGATGCTAAGGTAAGCCAGCTCATAC 5460  
Qy 5461 CCACCATAGGTCCTAGAGTCTAGAGCTGCAGTCAGCTCACTCACTGAGGTCGCAAGATGC 5520  
Db 5461 CCACCATAGGTCCTAGAGTCTAGAGCTGCAGTCAGCTCACTCACTGAGGTCGCAAGATGC 5520  
Qy 5521 CTCTAAAGATGTAGGAAAGTGTAGAGAGGGGTGAGGTTGCGGCTCCGGGTGAGAGTG 5580  
Db 5521 CTCTAAAGATGTAGGAAAGTGTAGAGAGGGGTGAGGTTGCGGCTCCGGGTGAGAGTG 5580  
Qy 5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCAATTTGGCTTTGGGAACTGCAGTCTCTCT 5640  
Db 5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCAATTTGGCTTTGGGAACTGCAGTCTCTCT 5640  
Qy 5641 GGGGAGCTGATGTGAATGATCTTGGGTGGATCC 5674  
Db 5641 GGGGAGCTGATGTGAATGATCTTGGGTGGATCC 5674

## RESULT 3

Q98902  
ID Q98902 standard; DNA: 5724 BP.  
AC Q98902:  
DT 28-FEB-1996 (first entry)  
DE Tumour rejection antigen (MAGE-1) gene.  
KW Tumour rejection antigen; MAGE-1; monoclonal antibody; Mab;  
KW diagnosis; immunoassay; cancer; ss.  
OS Homo sapiens.  
FH key Location/Qualifiers  
FT cds 3881..4711  
FT /\*tag= a  
FT /product= Tumour rejection antigen MAGE-1.  
FT /note= "The CDS is not indicated in the text of the  
FT specification but is suggested in the layout  
FT of the sequence."  
PN W09520974-A1.  
PD 10-AUG-1995.  
PF 05-JAN-1995; U00095.  
PR 01-FEB-1994; US-190411.  
PA (LUDW-) LUDWIG INST CANCER RES.  
PA (SLOK) SLOAN KETTERING INST CANCER RES.

PA (SLOK) MEMORIAL SLOAN-KETTERING CANCER CENT.  
PI Boon-fallour T, Chen Y, Garin-chesa P, Old LJ, Rettig WJ;  
PI Stockert E, Van der bruggen P;  
DR WPI: 95-283606/37.  
PT New monoclonal antibody binding specifically to MAGE-1 - useful for  
PT diagnosis and monitoring of cancer, also new hybridomas, recombinant  
PT MAGE-1 and immunogenic peptide(s)  
PS Disclosure: Page 16-19; 33pp; English.  
CC A monoclonal antibody directed against the tumour rejection antigen  
CC (MAGE-1) can be used to detect MAGE-1 in samples by standard  
CC immunassay methods for diagnosis and monitoring of cancer etc. The  
CC monoclonal antibody is designated MA454 and is produced by the  
CC hybridoma deposited as ATCC HB11540. The monoclonal antibody is  
CC specific for MAGE-1, having no reactivity for MAGE-2 or MAGE-3.  
CC Peptide fragments of MAGE-1 (See R0618-20) may be useful as  
CC immunogens for production of the monoclonal antibody and antisera.  
SQ Sequence 5724 BP; 1282 A; 1653 C; 1589 G; 1200 T;

Query Match 99.6%; Score 5650.8; DB 1; Length 5724;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5652; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCGGGGCACCACCTGGCATCCCTCCCTCCCTACACCCCAATCCCTCCCTTTAGCCACCC 60  
Db 1 CCCGGGGCACCACCTGGCATCCCTCCCTCCCTACACCCCAATCCCTCCCTTTAGCCACCC 60  
Qy 61 ATCCAAACATCTTCACGCTCACGCCAGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGG 120  
Db 61 ATCCAAACATCTTCACGCTCACGCCAGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGG 120  
Qy 121 CTCTCAACCCAGGGAAGCCAGGTCGCCAGATGTGACGCCACTGACTTGAGCATTTAGTGG 180  
Db 121 CTCTCAACCCAGGGAAGCCAGGTCGCCAGATGTGACGCCACTGACTTGAGCATTTAGTGG 180  
Qy 181 TTAGAGAGAGGAGGAGGTTTTCGGTCTGAGGGGCGGCTGAGATCGGTGAGGAGGAGG 240  
Db 181 TTAGAGAGAGGAGGAGGTTTTCGGTCTGAGGGGCGGCTGAGATCGGTGAGGAGGAGG 240  
Qy 241 CCACGCTCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
Db 241 CCACGCTCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
Qy 301 AGATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
Db 301 AGATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
Qy 361 TCAGGCTGGGCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCC 420  
Db 361 TCAGGCTGGGCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCC 420  
Qy 421 CTCCTGTGTATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
Db 421 CTCCTGTGTATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
Qy 481 CATGCTCAGGATTTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
Db 481 CATGCTCAGGATTTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
Qy 541 CCCCACTCAATGCTCACTCCCGTGACCCCAACCCCGTCTTTCATTTGTCATTCACACCCCA 600  
Db 541 CCCCACTCAATGCTCACTCCCGTGACCCCAACCCCGTCTTTCATTTGTCATTCACACCCCA 600  
Qy 601 CCCCACTCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCG 660  
Db 601 CCCCACTCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCG 660  
Qy 661 CACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCG 720  
Db 661 CACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCG 720  
Qy 721 CCAGGAACATCCGGGTGCCGGATGTGAGGCACTGACTTGGCATTTGGGGGAGAGA 780  
Db 721 CCAGGAACATCCGGGTGCCGGATGTGAGGCACTGACTTGGCATTTGGGGGAGAGA 780

Db 721 CCAGGAACATCCGGGTCCCGGATGTACGCCACTGACATTGCGCATTTGTGGGCAGAGA 780  
Qy 781 GAACCGAGGTTTCCATTCTAGGACGCGGTAGAGTTTCGGCCGAAGAACTGTACCCAGG 840  
Db 781 GAACCGAGGTTTCCATTCTAGGACGCGGTAGAGTTTCGGCCGAAGAACTGTACCCAGG 840  
Qy 841 CTCGTGTAGGAGGCAAGGTGAGAGGCTGAGGAGGACTGAGGACCCGCGCACTCAAATA 900  
Db 841 CTCGTGTAGGAGGCAAGGTGAGAGGCTGAGGAGGACTGAGGACCCGCGCACTCAAATA 900  
Qy 901 GAGAGCCCCAAATATTCAGCCCCGCGCTTGTCTGCCAGCCCTGCCCCACCCGCGGGAAGA 960  
Db 901 GAGAGCCCCAAATATTCAGCCCCGCGCTTGTCTGCCAGCCCTGCCCCACCCGCGGGAAGA 960  
Qy 961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAGGCTTTGAGAGACACCAAGTTTC 1020  
Db 961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAGGCTTTGAGAGACACCAAGTTTC 1020  
Qy 1021 TTTCTCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAAGGCGAGGACTGGTTAGGAGAGG 1080  
Db 1021 TTTCTCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAAGGCGAGGACTGGTTAGGAGAGG 1080  
Qy 1081 GCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCCAAGAGGGAGGCTGTGGGCC 1140  
Db 1081 GCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCCAAGAGGGAGGCTGTGGGCC 1140  
Qy 1141 CCAAGACTGCACCTCCAATCCCACTCCACCCCATTCGCAATTCCTCCATCCCAACCAACC 1200  
Db 1141 CCAAGACTGCACCTCCAATCCCACTCCACCCCATTCGCAATTCCTCCATCCCAACCAACC 1200  
Qy 1201 CCCATCTCCTCAGCTACACCTTCCACCCCATTCCTACTCTCTACTCTCCTACTCCTGACCAACC 1260  
Db 1201 CCCATCTCCTCAGCTACACCTTCCACCCCATTCCTACTCTCTACTCCTGACCAACC 1260  
Qy 1261 ACCCTCAGCCCCAGCACCCAGCCCAACCTTCTGCCACCTCACCCTCAGTGCCTCCCAAC 1320  
Db 1261 ACCCTCAGCCCCAGCACCCAGCCCAACCTTCTGCCACCTCACCCTCAGTGCCTCCCAAC 1320  
Qy 1321 CCCACCTCATCTCTCATGTGCCCACTCCCATCGCCTCCCACTTCTGGCAGAAATCC 1380  
Db 1321 CCCACCTCATCTCTCATGTGCCCACTCCCATCGCCTCCCACTTCTGGCAGAAATCC 1380  
Qy 1381 GGTTCGCTCTCTCAACCCAGGAGGAGCCCTGGTAGGCCGATGTGAACCACTGACT 1440  
Db 1381 GGTTCGCTCTCTCAACCCAGGAGGAGCCCTGGTAGGCCGATGTGAACCACTGACT 1440  
Qy 1441 TGAACCTCAGAGATCTGAGAGAACCCAGGTTCAATTTAATGTTCTGAGGGCGGCTTGAG 1500  
Db 1441 TGAACCTCAGAGATCTGAGAGAACCCAGGTTCAATTTAATGTTCTGAGGGCGGCTTGAG 1500  
Qy 1501 ATCCACTCAGGGAGTGGTTTATGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGG 1560  
Db 1501 ATCCACTCAGGGAGTGGTTTATGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGG 1560  
Qy 1561 ACTGAGGAGGCACACCCAGGTAGATGGCCCCAAATGATCAGTACCAACCCCTGCTG 1620  
Db 1561 ACTGAGGAGGCACACCCAGGTAGATGGCCCCAAATGATCAGTACCAACCCCTGCTG 1620  
Qy 1621 CCAGCCCTGGACCAACCCGCGCAGCAGATGCTCAGCTGGACCAACCCCGCTCCGCTCC 1680  
Db 1621 CCAGCCCTGGACCAACCCGCGCAGCAGATGCTCAGCTGGACCAACCCCGCTCCGCTCC 1680  
Qy 1681 CACTGCCCATTAAACCCACAGGCAATCTGTAGTCATAGCTTATGTACCCGGGCGAGGTT 1740  
Db 1681 CACTGCCCATTAAACCCACAGGCAATCTGTAGTCATAGCTTATGTACCCGGGCGAGGTT 1740  
Qy 1741 GGTGAGGAGGAGGCGCCAGGCATCAAGGTCCAGCATCCGCCCCGCAATTAGGGTCAGG 1800  
Db 1741 GGTGAGGAGGAGGCGCCAGGCATCAAGGTCCAGCATCCGCCCCGCAATTAGGGTCAGG 1800  
Qy 1801 ACCCTGGGAGGAACTGAGGTTTCCCAACCCACACCTGTCTCTCATCTCCACCCGCAACC 1860  
Db 1801 ACCCTGGGAGGAACTGAGGTTTCCCAACCCACACCTGTCTCTCATCTCCACCCGCAACC 1860

Qy 1861 CCACCTACATTTCCCATACCTACCCCTACCCCAACCTCATCTTGTGAGAACTCCCTGCTG 1920  
Db 1861 CCACCTACATTTCCCATACCTACCCCTACCCCAACCTCATCTTGTGAGAACTCCCTGCTG 1920  
Qy 1921 TCACCCACCGAAGCCACGGGAATFGCGGCCACGACATCGGATTTGACGTTCCTCATCCA 1980  
Db 1921 TCACCCACCGAAGCCACGGGAATFGCGGCCACGACATCGGATTTGACGTTCCTCATCCA 1980  
Qy 1981 GGGTCTCATGAGGGAAGGGCTTGAACAGGGCTCTAGGGAGCAGAGGGAGGCGCTAC 2040  
Db 1981 GGGTCTCATGAGGGAAGGGCTTGAACAGGGCTCTAGGGAGCAGAGGGAGGCGCTAC 2040  
Qy 2041 TGCAGATGAGGAGGCGCTCAGAGGACCCAGCACCTTAGACACCCGACCCCTCTCTGAG 2100  
Db 2041 TGCAGATGAGGAGGCGCTCAGAGGACCCAGCACCTTAGACACCCGACCCCTCTCTGAG 2100  
Qy 2101 ACTGAGGCTGCCACTTCTGCGCTCAAGAAATCAGAAACGATGGGACTCAGATTGCGATGGG 2160  
Db 2101 ACTGAGGCTGCCACTTCTGCGCTCAAGAAATCAGAAACGATGGGACTCAGATTGCGATGGG 2160  
Qy 2161 GTGGGACCCAGGCTGCAAGGCTTACCGGAGGAAGAGGAGGAGGACTCAGGGGACCTT 2220  
Db 2161 GTGGGACCCAGGCTGCAAGGCTTACCGGAGGAAGAGGAGGAGGACTCAGGGGACCTT 2220  
Qy 2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGCTGCGCCACATATG 2280  
Db 2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGCTGCGCCACATATG 2280  
Qy 2281 GCCCATATTTCTGTCATCTTTGAGGTTGACAGGACAGAGCTGTGCTGTGAGAGTGGGGCC 2340  
Db 2281 GCCCATATTTCTGTCATCTTTGAGGTTGACAGGACAGAGCTGTGCTGTGAGAGTGGGGCC 2340  
Qy 2341 TCAGGTCAACAGAGGAGGAGTTCCAGGATCCATATGSCCAAGATGTCGCCCTTCATG 2400  
Db 2341 TCAGGTCAACAGAGGAGGAGTTCCAGGATCCATATGSCCAAGATGTCGCCCTTCATG 2400  
Qy 2401 AGGACTGGGATATCCCGGCTCAGAAAGAGGACTTCCACAGTCTGGCTGTGCCCTT 2460  
Db 2401 AGGACTGGGATATCCCGGCTCAGAAAGAGGACTTCCACAGTCTGGCTGTGCCCTT 2460  
Qy 2461 TTAGTAGCTCTAGGGGACCAAGATCAGGGATGGGGTATGTTCCATTTCTCATTGTACCA 2520  
Db 2461 TTAGTAGCTCTAGGGGACCAAGATCAGGGATGGGGTATGTTCCATTTCTCATTGTACCA 2520  
Qy 2521 CAGCAGGAAGTTGGGGGCGCTCAGGAGATGGGTCTTGGGTAAGGGGGATGCT 2580  
Db 2521 CAGCAGGAAGTTGGGGGCGCTCAGGAGATGGGTCTTGGGTAAGGGGGATGCT 2580  
Qy 2581 ACTCATGTCAGGGAATTTGGGGTTGAGGAAGCACAGCGCTGGCAGGAATAAGATGAGT 2640  
Db 2581 ACTCATGTCAGGGAATTTGGGGTTGAGGAAGCACAGCGCTGGCAGGAATAAGATGAGT 2640  
Qy 2641 GAGACAGCAAGGCTATTGGAATCCACCCAGAACCCAAAGGGGTGAGCCCTGGACACC 2700  
Db 2641 GAGACAGCAAGGCTATTGGAATCCACCCAGAACCCAAAGGGGTGAGCCCTGGACACC 2700  
Qy 2701 TCACCCAGGATGGCTTCTTTTCTACTCTCTTCCAGATCTGGGCGAGGTGAGGACCT 2760  
Db 2701 TCACCCAGGATGGCTTCTTTTCTACTCTCTTCCAGATCTGGGCGAGGTGAGGACCT 2760  
Qy 2761 CATTTCTCAGAGGTTGACTCAGGTCAAGTACCTAGGACCCCTCTGTTCTAAGACAGAGCG 2820  
Db 2761 CATTTCTCAGAGGTTGACTCAGGTCAAGTACCTAGGACCCCTCTGTTCTAAGACAGAGCG 2820  
Qy 2821 GTCCAGGATCTGCCATGCGTTCGGGTGAGAAACATGAGGAGGACTGAGGGTACCCAG 2880  
Db 2821 GTCCAGGATCTGCCATGCGTTCGGGTGAGAAACATGAGGAGGACTGAGGGTACCCAG 2880  
Qy 2881 GACCAGAACACTGAGGAGGACTGCACAGAAATCAGCCCTGCTCCCTGCTACCCCGAGAG 2940  
Db 2881 GACCAGAACACTGAGGAGGACTGCACAGAAATCAGCCCTGCTCCCTGCTACCCCGAGAG 2940



Db	5101	T	A	G	A	G	T	T	G	T	G	T	T	T	A	T	T	C	A	G	A	T	T	G	G	A	A	T	C	A	T	T	T	T	G	A	A	T	T	G	G	A	A	T		5160							
Qy	5161	A	A	C	A	G	A	G	T	G	G	A	A	T	A	G	T	A	C	T	T	A	G	A	A	T	G	T	A	A	A	A	T	G	A	A	A	A	T	G	A	A	A	A	T		5220						
Db	5161	A	C	A	G	C	A	G	T	G	G	A	A	T	A	G	T	A	C	T	T	A	G	A	A	T	G	T	A	A	A	A	T	G	A	A	A	A	T	G	A	A	A	T		5220							
Qy	5221	A	G	A	C	T	A	A	G	A	A	A	T	T	A	A	G	A	T	A	G	T	A	C	T	T	T	G	C	T	T	A	T	T	G	C	T	T	A	T	T	C	T	G	A	A		5280					
Db	5221	A	G	A	C	T	A	A	G	A	A	A	T	T	A	A	G	A	T	A	G	T	A	C	T	T	T	G	C	T	T	A	T	T	G	C	T	T	A	T	T	C	T	G	A	A		5280					
Qy	5281	A	T	T	T	T	A	A	G	A	T	A	T	A	T	A	C	A	T	A	C	T	G	G	A	T	T	C	T	T	T	G	G	C	T	T	T	T	G	A	G	A	T	G	T	A	A		5340				
Db	5281	A	T	T	T	T	A	A	G	A	T	A	T	A	T	A	C	A	T	A	C	T	G	G	A	T	T	C	T	T	T	G	G	C	T	T	T	T	G	A	G	A	T	G	T	A	A		5340				
Qy	5341	T	A	A	T	C	T	G	A	T	A	A	G	A	A	T	T	C	T	T	C	A	C	T	G	G	C	T	T	T	T	T	C	A	C	T	G	G	C	T	T	T	T	C	A	C	T	G	A	C	A		5400
Db	5341	T	A	A	T	C	T	G	A	T	A	A	G	A	A	T	T	C	T	T	C	A	C	T	G	G	C	T	T	T	T	T	C	A	C	T	G	G	C	T	T	T	T	C	A	C	T	G	A	C	A		5400
Qy	5401	T	C	T	G	C	T	T	T	T	G	A	G	C	C	C	T	G	G	T	T	A	G	T	A	G	T	G	G	A	T	G	C	T	A	A	G	T	A	A	G	C	C	A	G	A	C	T	A	T		5460	
Db	5401	T	C	T	G	C	T	T	T	T	G	A	G	C	C	C	T	G	G	T	T	A	G	T	A	G	T	G	G	A	T	G	C	T	A	A	G	T	A	A	G	C	C	A	G	A	C	T	A	T		5460	
Qy	5461	C	C	A	C	C	A	T	A	G	G	T	C	G	T	A	G	A	G	T	A	C	A	G	T	C	A	G	T	A	C	A	G	T	A	C	A	G	T	A	C	A	G	T	A	C	A	G	T		5520		
Db	5461	C	C	A	C	C	A	T	A	G	G	T	C	G	T	A	G	A	G	T	A	C	A	G	T	C	A	G	T	A	C	A	G	T	A	C	A	G	T	A	C	A	G	T	A	C	A	G	T		5520		
Qy	5521	C	T	C	T	A	A	G	A	T	G	T	A	G	G	A	A	A	A	A	G	T	A	G	A	G	G	G	T	A	G	G	T	A	G	G	C	T	C	C	G	G	T	C	A	G	A	G	T		5580		
Db	5521	C	T	C	T	A	A	G	A	T	G	T	A	G	G	A	A	A	A	A	G	T	A	G	A	G	G	G	T	A	G	G	T	A	G	G	C	T	C	C	G	G	T	C	A	G	A	G	T		5580		
Qy	5581	G	T	G	A	G	T	G	T	C	A	A	T	G	C	C	C																																				

RESULT	4	
Q32351		
ID	Q32351 standard; DNA; 2419 BP.	
AC	Q32351;	
DT	22-APR-1993 (first entry)	
DE	Antigen E gene.	
KE	Stable; antigen; E; D; F; A; human; melanoma; cell line; MZ2-MEL;	
KW	cytolytic T cell; MEL3.1; open reading frame; homology; MAGE;	
KW	melanoma antigen; ss.	
OS	Homo sapiens.	
PN	W09220356-A.	
PD	26-NOV-1992.	
PF	22-MAY-1992; U04354.	
PR	23-MAY-1991; US-705702.	
PR	09-JUL-1991; US-728838.	
PR	23-SEP-1991; US-764364.	
PR	12-DEC-1991; US-807043.	
PA	(LUDW-) LUDWIG INST CANCER RES.	
PI	Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;	
PI	Van den Eynde B, Van Bruggen P, Van Pel A;	
DR	NPL; 92-415460/50.	
PT	Neuric acid mol. encoding a human tumour rejection antigen	
PT	precursor - useful as an immunostimulant in a vaccine for	
PT	treating and preventing cancers, also useful in diagnosis .	
PS	Disclosure; Page 69-70; 14pp; English.	
CC	This sequence encodes the stable antigen E. This antigen is expressed	
CC	along with antigens "D, F and A" by the human melanoma cell line MZ2-	
CC	MEL. These antigens are all recognised by cytolytic T cells. A	
CC	CC subline of MZ2-MEL is MEL3.1 which only expresses antigen E. This	
CC	cell line was chosen as a source for the isolation of this sequence.	
CC	This sequence was found to contain three exons. The open reading frame	
CC	for antigen E was contained within the first two exons. An ATG is	
CC	located at position 66 of exon 3 and is followed by an 828 base pair	
CC	reading frame. The three exons contain 65, 73 and 1551 base pairs.	



Qy	4156	TATCTGGAGTCTCTGTTCCGAGCAGTAACTACCTAAGAAGTGCTGCTGATTTGGTTGGTTT	4211
Db	901	TATCTCGGAGTCTCTGTTCCGAGCAGTAACTACCTAAGAAGTGCTGCTGATTTGGTTGGTTT	960
Qy	4216	TCCTGCTCCTCAATATCATGAGCCAGGAGCCAGTCACAAAGCACAANAATGCTGGAGAGTGT	4275
Db	961	TCCTGCTCCTCAATATCATGAGCCAGGAGCCAGTCACAAAGCACAANAATGCTGGAGAGTGT	1020
Qy	4276	CATCAAAAATATCAAGCACTGTTTTCTTGAGATCTTCGGCAAAAGCCCTCTGAGTCCCTTGCA	4335
Db	1021	CATCAAAAATATCAAGCACTGTTTTCTTGAGATCTTCGGCAAAAGCCCTCTGAGTCCCTTGCA	1080
Qy	4336	GCTGGTCTTTGGCATTTGAGCTGAAGAGACAGACCCACCGGCCACATCTCTATGTCTCTGT	4395
Db	1081	GCTGGTCTTTGGCATTTGAGCTGAAGAGAGACAGACCCACCGGCCACATCTCTATGTCTCTGT	1140
Qy	4396	CACCTGCCCTAGGTCTCTCTATGATGGCTGCTGGTGATAAATCAGATCATGCCCAAGAC	4455
Db	1141	CACCTGCCCTAGGTCTCTCTATGATGGCTGCTGGTGATAAATCAGATCATGCCCAAGAC	1200
Qy	4456	AGGCTCTCTGATAAATGTCTCGTCAATGATTCGAATGAGGCGGCCCATGCTCCTCAGGA	4515
Db	1201	AGGCTCTCTGATAAATGTCTCGTCAATGATTCGAATGAGGCGGCCCATGCTCCTCAGGA	1260
Qy	4516	GGAATCTCGGAGGAGCTGAGTGTGATGGAGTGATGATGGGAGGAGACACAGTGCCTA	4575
Db	1261	GGAATCTCGGAGGAGCTGAGTGTGATGGAGTGATGATGGGAGGAGACACAGTGCCTA	1320
Qy	4576	TGGGAGGCCAGGAGCTGCTCACCCAGATTTGGTGAGGAAAAGTACCTGGAGTACGG	4635
Db	1321	TGGGAGGCCAGGAGCTGCTCACCCAGATTTGGTGAGGAAAAGTACCTGGAGTACGG	1380
Qy	4636	CAGGTGCCGGACAGTGATCCCGACGCTATGAGTTCCTTGCTGGGTCCAAGGGCCCTCGCT	4695
Db	1381	CAGGTGCCGGACAGTGATCCCGACGCTATGAGTTCCTTGCTGGGTCCAAGGGCCCTCGCT	1440
Qy	4696	GAAACAGCTATGTGAAGTCTCTGAGTATGTGATCAAGTTCAGTGAAGTTCGCTTT	4755
Db	1441	GAAACAGCTATGTGAAGTCTCTGAGTATGTGATCAAGTTCAGTGAAGTTCGCTTT	1500
Qy	4756	TTCTTCCCATCCCTGGTCAACACGCTTTGAGAGAGGAGGAAGGAGGTCTGAGCATGA	4815
Db	1501	TTCTTCCCATCCCTGGTCAACACGCTTTGAGAGAGGAGGAAGGAGGTCTGAGCATGA	1560
Qy	4816	GTTGCAGCCAAGGCCAGTGGGAGGGGAGCTGGGCCAGTGCACCTTCCAGGGCCGCGTCCA	4875
Db	1561	GTTGCAGCCAAGGCCAGTGGGAGGGGAGCTGGGCCAGTGCACCTTCCAGGGCCGCGTCCA	1620
Qy	4876	GCAGCTTCCCTTGCCTGTGTGACATGAGGCCCATTTCTTCACTCTGAAGAGAGCGGTTCAG	4935
Db	1621	GCAGCTTCCCTTGCCTGTGTGACATGAGGCCCATTTCTTCACTCTGAAGAGAGCGGTTCAG	1680
Qy	4936	TGTTCTCAGTAGTAGTTCCTCTTATTTGGGTGACTTGGAGATTTATCTTTGTTCTCTTT	4995
Db	1681	TGTTCTCAGTAGTAGTTCCTCTTATTTGGGTGACTTGGAGATTTATCTTTGTTCTCTTT	1740
Qy	4996	TTGGAATTTGTTCAAAATGTTTTTTTTTAAGGGATGGTTTGAATGAACCTTCAGCATCCAGTT	5055
Db	1741	TTGGAATTTGTTCAAAATGTTTTTTTTTAAGGGATGGTTTGAATGAACCTTCAGCATCCAGTT	1800
Qy	5056	TATGAATGACAGCGTACACAGTTCCTGTGTATATAGTTTTAAGGGTAAAGAGTCTTGTGTT	5115
Db	1801	TATGAATGACAGCGTACACAGTTCCTGTGTATATAGTTTTAAGGGTAAAGAGTCTTGTGTT	1860
Qy	5116	TTATTAGATTTGGGAATCCATCTATTTTTTGTGAATTCGGGATATAACACGAGTGGGAATA	5175
Db	1861	TTATTAGATTTGGGAATCCATCTATTTTTTGTGAATTCGGGATATAACACGAGTGGGAATA	1920
Qy	5176	AGTACTTTAGAAATGTGAAAATGACAGTAAATAGATGACATAAAGAACTTAAGAAATTT	5235
Db	1921	AGTACTTTAGAAATGTGAAAATGACAGTAAATAGATGACATAAAGAACTTAAGAAATTT	1980
Qy	5236	AAGAGATAGTCAATTTCTGCCCTTATACCTCAGTCTATTCTCTAAAAATTTTTTAAGATATA	5295

Db	1981	AAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTCTAAAAATTTTAAAGATATA	2040
Qy	5296	TGCATACCTGGATTTCCTTTGGCTTCTTTTGACAATGTAAGAGAAAATTAATCTGAATAAAG	5355
Db	2041	TGCATACCTGGATTTCCTTTGGCTTCTTTTGAGATGTAAGAGAAATTAATCTGAATAAAG	2100
Qy	5356	AATTCTTCCTGTTTCACATGGCTCTTTTCTCTCCATGCACCTGAGCATCTGCTTTTGGGAAG	5415
Db	2101	AATTCTTCCTGTTTCACATGGCTCTTTTCTCTCCATGCACCTGAGCATCTGCTTTTGGGAAG	2160
Qy	5416	GCCCTGGGTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCACCATAGGTCG	5475
Db	2161	GCCCTGGGTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCACCATAGGTCG	2220
Qy	5476	TAGAGTCTAGGAGCTGCAGTCAACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAGG	5535
Db	2221	TAGAGTCTAGGAGCTGCAGTCAACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAGG	2280
Qy	5536	GAAGAAGTAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTGGTGAGTGTCAATGC	5595
Db	2281	GAAGAAGTAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTGGTGAGTGTCAATGC	2340
Qy	5596	CCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTCTCGGGGAGCTCATTTGT	5655
Db	2341	CCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTCTCGGGGAGCTCATTTGT	2400
Qy	5656	AATGATCTTGGGTGGATCC	5674
Db	2401	AATGATCTTGGGTGGATCC	2419

RESULT 5

Q72476

ID Q72476 standard; DNA; 2419 BP.

AC Q72476;

DT 21-JUN-1995 (first entry)

DE Tumour rejection antigen E encoding DNA.

KW Tumour rejection antigen E; melanoma antigen-3; MAGE-3;

XW cancer; cytolytic T cells; antigen D; human leucocyte antigen; ss.

OS Homo sapiens.

PN W09423031-A.

PD 13-OCT-1994.

PF 17-MAR-1994; U02877.

PR 26-MAR-1993; US-037230.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Boon-Falloux T, Gauglier B, Van DEN EYNDE B, Van DER BRUGGEN P;

DR WPI; 94-333192/41.

PT New tumour rejection antigen precursor MAGE3 - useful in treatment and diagnosis of cancer

PT Disclosure; Page 58; 105pp; English.

PS Q72476 encodes tumour rejection antigen E, another sequence

CC Q72476 encodes melanoma antigen-3 (MAGE-3) a tumour rejection

CC antigen precursor. Melanomas characterised by the expression of

CC MAGE-3 can be detected, or monitored, by contacting a test sample

CC with an agent that can recognise MAGE-3. The melanoma can be treated

CC by the administration of cytolytic T cells specific for the complex of

CC antigen D (the mature rejection antigen derived from MAGE-3) and a human

CC leucocyte antigen (esp. HLA-A1).

CC Sequence 2419 BP; 562 A; 581 C; 677 G; 599 T;

SQ

	Query Match	42.6%	Score 2419;	DB 1;	Length 2419;	
	Best Local Similarity	100.0%;	Pred. No. 0;			
	Matches 2419;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	3256	GGATCCAGGCCCCGCCAGGAAATAATAAGGGCCCTGCGTGAGACACAGAGGGGTCA	TCCC	3315		
Dd	1	GGATCCAGGCCCCGCCAGGAAATAATAAGGGCCCTGCGTGAGACACAGAGGGGTCA	TCCC	60		
Qy	3316	ACTGCATCAGAGTGGGGATGTTCACAGATCCAGCCACCCTCTCGGTAGCACTGAGA	AGC	3375		
Dd	61	ACTGCATCAGAGTGGGGATGTTCACAGATCCAGCCACCCTCTCGGTAGCACTGAGA	AGC	120		



QY 3376 CAGGGCTGTGCTTGGGCTGTGACACCTGAGGCGCCGGATTCCTCTTCTGGAGCTCCA 3435  
DB 121 CAGGGCTGTGCTTGGGCTGTGACACCTGAGGCGCCGGATTCCTCTTCTGGAGCTCCA 180  
QY 3436 GGAACACAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAACAGACAGAGGATG 3495  
DB 181 GGAACACAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAACAGACAGAGGATG 240  
QY 3496 CACAGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAACAGAGGATG 3555  
DB 241 CACAGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAACAGAGGATG 300  
QY 3556 CAGGACACATAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAACAGAGGATG 3615  
DB 301 CAGGACACATAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAACAGAGGATG 360  
QY 3616 CGACCTCTGTGGGCGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAAC 3675  
DB 361 CGACCTCTGTGGGCGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAAC 420  
QY 3676 GGAACAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAAC 3735  
DB 421 GGAACAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAAC 480  
QY 3736 GATCTGTAAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAAC 3795  
DB 481 GATCTGTAAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCTCAGGTCAAC 540  
QY 3796 CACACTCTCTCTCCGAGGCTGTGGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCT 3855  
DB 541 CACACTCTCTCTCCGAGGCTGTGGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCT 600  
QY 3856 GCCTGTGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCT 3915  
DB 601 GCCTGTGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGACAGTATCT 660  
QY 3916 TGAGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAG 3975  
DB 661 TGAGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAG 720  
QY 3976 CTCCT 4035  
DB 721 CTCCT 780  
QY 4036 AGATCT 4095  
DB 781 AGATCT 840  
QY 4096 ACAGAGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTG 4155  
DB 841 ACAGAGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTG 900  
QY 4156 TATCT 4215  
DB 901 TATCT 960  
QY 4216 TCT 4275  
DB 961 TCT 1020  
QY 4276 CATCAAAATTTACAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGG 4335  
DB 1021 CATCAAAATTTACAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGG 1080  
QY 4336 GCTGGTCTTTGGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGT 4395  
DB 1081 GCTGGTCTTTGGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGT 1140  
QY 4396 CACCT 4455  
DB 1141 CACCT 1200

QY 4456 AGGCTTCTCTGATAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4515  
DB 1201 AGGCTTCTCTGATAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
QY 4516 GGAATCTGGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTG 4575  
DB 1261 GGAATCTGGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTG 1320  
QY 4576 TGGGAGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTG 4635  
DB 1321 TGGGAGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTG 1380  
QY 4636 CAGGTGCTGGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTG 4695  
DB 1381 CAGGTGCTGGGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCT 1440  
QY 4696 GAAACACAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGG 4755  
DB 1441 GAAACACAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGG 1500  
QY 4756 TTCTTCT 4815  
DB 1501 TTCTTCT 1560  
QY 4816 GTTCT 4875  
DB 1561 GTTCT 1620  
QY 4876 GCAGCTTCT 4935  
DB 1621 GCAGCTTCT 1680  
QY 4936 TGTCT 4995  
DB 1681 TGTCT 1740  
QY 4996 TTGGAATTTCT 5055  
DB 1741 TTGGAATTTCT 1800  
QY 5056 TATGAATGACAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTG 5115  
DB 1801 TATGAATGACAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTG 1860  
QY 5116 TTATTCAGATTTGGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGG 5175  
DB 1861 TTATTCAGATTTGGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGG 1920  
QY 5176 AGTACTTCT 5235  
DB 1921 AGTACTTCT 1980  
QY 5236 AAGAGATGATCAATTTCT 5295  
DB 1981 AAGAGATGATCAATTTCT 2040  
QY 5296 TGCATCT 5355  
DB 2041 TGCATCT 2100  
QY 5356 AATTCCT 5415  
DB 2101 AATTCCT 2160  
QY 5416 GCGCTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGG 5475  
DB 2161 GCGCTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGG 2220  
QY 5476 TAGAGTCTAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGT 5535  
DB 2221 TAGAGTCTAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGT 2280  
QY 5536 GAAAGTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGTGAGGAGTGTGGGCTGT 5595



Qy 4756 TTCTTCCATCCCTCGCTGAAGCAGCTTTGAGAGAGGAGGAGGAGTCTGAGCATGA 4815  
Db 1501 TTCTTCCATCCCTCGCTGAAGCAGCTTTGAGAGAGGAGGAGGAGTCTGAGCATGA 1560  
Qy 4816 GTTCAGCAAGGCGAGTGGGAGGGGAGTGGGCGAGTGCACCTTTCCAGGGCCGCTCCA 4875  
Db 1561 GTTCAGCAAGGCGAGTGGGAGGGGAGTGGGCGAGTGCACCTTTCCAGGGCCGCTCCA 1620  
Qy 4876 GCAGCTCCCTCGCTCGTGTGACATGAGGCCCATCTTCTACTCTGAAGAGCGGTCAG 4935  
Db 1621 GCAGCTCCCTCGCTCGTGTGACATGAGGCCCATCTTCTACTCTGAAGAGCGGTCAG 1680  
Qy 4936 TGTCTCAGTAGTAGTCTCTCTATTTGGGTGACCTTGGAGATTTATCTTTGTTCTCTT 4995  
Db 1681 TGTCTCAGTAGTAGTCTCTCTATTTGGGTGACCTTGGAGATTTATCTTTGTTCTCTT 1740  
Qy 4996 TTGGAATGTTCAATGTTTTTTTTTAAAGGATGGTTGAATGAATCACTTCAGCATCAAAGTT 5055  
Db 1741 TTGGAATGTTCAATGTTTTTTTTTAAAGGATGGTTGAATGAATCACTTCAGCATCAAAGTT 1800  
Qy 5056 TATCAATGACAGCAGTGCACAGTCTCTGTATATAGTTTAAAGGTAGAGTCTTGTCTT 5115  
Db 1801 TATCAATGACAGCAGTGCACAGTCTCTGTATATAGTTTAAAGGTAGAGTCTTGTCTT 1860  
Qy 5116 TTATTAGATTGGGAATCCATCTATTTTGTGAATGGGATAATAACAGCAGTGAATA 5175  
Db 1861 TTATTAGATTGGGAATCCATCTATTTTGTGAATGGGATAATAACAGCAGTGAATA 1920  
Qy 5176 AGTACTTAGAATGTGAATAGAGCAGTAAATAGATGAGATGAAGAACTAAAGAAAT 5235  
Db 1921 AGTACTTAGAATGTGAATAGAGCAGTAAATAGATGAGATGAAGAACTAAAGAAAT 1980  
Qy 5236 AAGAGATAGTCAATCTTGCCTTATACCTCAGTCTATCTGTAAATTTTAAAGATATA 5295  
Db 1981 AAGAGATAGTCAATCTTGCCTTATACCTCAGTCTATCTGTAAATTTTAAAGATATA 2040  
Qy 5296 TGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAAGAAATTAATCTGAATAAG 5355  
Db 2041 TGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAAGAAATTAATCTGAATAAG 2100  
Qy 5356 AATCTTCTCTTCACTGGCTTTCTTCTTCCATGCACTGAGCATCTGCTTTTGGAG 5415  
Db 2101 AATCTTCTCTTCACTGGCTTTCTTCTTCCATGCACTGAGCATCTGCTTTTGGAG 2160  
Qy 5416 GCCCTGGTGTAGTAGTGGAGATGTAAGGTAAAGCAGACTCATACCCACCATAGGTCG 5475  
Db 2161 GCCCTGGTGTAGTAGTGGAGATGTAAGGTAAAGCAGACTCATACCCACCATAGGTCG 2220  
Qy 5476 TAGAGTCTAGGAGCTGCAGTCACTGAGTGGGCAAGATGCTCTTAAAGATGTAGG 5535  
Db 2221 TAGAGTCTAGGAGCTGCAGTCACTGAGTGGGCAAGATGCTCTTAAAGATGTAGG 2280  
Qy 5536 GAAAGTGTAGAGAGGGTGGGGTGTGGGGTCCGGGTGAGAGTGGTGGAGTGTCAATGC 5595  
Db 2281 GAAAGTGTAGAGAGGGTGGGGTGTGGGGTCCGGGTGAGAGTGGTGGAGTGTCAATGC 2340  
Qy 5596 CCTGAGTGGGGATTTTGGGCTTTGGAACTCAGTCTCTTCTGGGGAGCTGATGT 5655  
Db 2341 CCTGAGTGGGGATTTTGGGCTTTGGAACTCAGTCTCTTCTGGGGAGCTGATGT 2400  
Qy 5656 AATGATCTTGGGTGGATCC 5674  
Db 2401 AATGATCTTGGGTGGATCC 2419

## RESULT 7

Q72472  
ID Q72472 standard; DNA: 2420 BP.

AC Q72472:

DT 21-JUN-1995 (first entry)

DE Tumour rejection antigen E precursor gene DNA.

KW Tumour antigen rejection precursor E; melanoma antigen-3; MAGE-3;

KW cancer; cytolytic T cells; antigen D; human leucocyte antigen; ss.  
OS Homo sapiens.  
PN WO9423031-A.  
PD 13-OCT-1994.  
PF 17-MAR-1994; U02877.  
PR 26-MAR-1993; US-037230.  
PA (LUDW-) LUDWIG INST CANCER RES.  
PI Boon-falleur T; Gaugier B, Van DEN EYNDE B, Van DER BRUGGEN P;  
DR WPI: 94-333192/41.  
PT New tumour rejection antigen precursor MAGE3 - useful in  
PT treatment and diagnosis of cancer  
PS Example 20; Page 28; 105pp; English.  
CC Q72472 is the tumour rejection antigen E precursor gene, another  
CC gene Q72470 encodes melanoma antigen-3 (MAGE-3) also a tumour rejection  
CC antigen precursor. Melanomas characterised by the expression of MAGE-3  
CC can be detected, or monitored, by contacting a test sample with an  
CC agent that can recognise MAGE-3. The melanoma can be treated by the  
CC administration of cytolytic T cells specific for the complex of antigen  
CC D (the mature rejection antigen derived from MAGE-3) and a human  
CC leucocyte antigen (esp. HLA-A1).  
SQ Sequence 2420 BP; 562 A; 582 C; 677 G; 599 T;

Query Match 42.4%; Score 2408; DB 1; Length 2420;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2419; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 3256 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTCGGTGAGAACAGAGGGGGTCTATCC 3315  
Db 1 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTCGGTGAGAACAGAGGGGGTCTATCC 60  
Qy 3316 ACTCATGAGAGTGGGGATGTACAGAGTCCAGGCCCTCTCTGGTAGCACTGAGAAGC 3375  
Db 61 ACTCATGAGAGTGGGGATGTACAGAGTCCAGGCCCTCTCTGGTAGCACTGAGAAGC 120  
Qy 3376 CAGGGCTGTCTTGGCTGTCTGCACCCCTGAGGGCCCGTGGATTCCTCTCTGGAGCTCCA 3435  
Db 121 CAGGGCTGTCTTGGCTGTCTGCACCCCTGAGGGCCCGTGGATTCCTCTCTGGAGCTCCA 180  
Qy 3436 GGAACACAGGAGTGGGGCTTGGTCTGAGACAGTATCTCAGGTACAGAGCAGAGATG 3495  
Db 181 GGAACACAGGAGTGGGGCTTGGTCTGAGACAGTATCTCAGGTACAGAGCAGAGATG 240  
Qy 3496 CACAGGGTGTGCCAGCAGTGAATGTTTCCCTGAATGCACCAAGGCCCTCCTGCCA 3555  
Db 241 CACAGGGTGTGCCAGCAGTGAATGTTTCCCTGAATGCACCAAGGCCCTCCTGCCA 360  
Qy 3556 CAGGACACATAGGACTCCAGAGTCTGGGCTCACCCTCTACTGTCTAGTCTCTAGAA 3615  
Db 301 CAGGACACATAGGACTCCAGAGTCTGGGCTCACCCTCTACTGTCTAGTCTCTAGAA 360  
Qy 3616 CGACCTCTGTGGCGGCTGTACCTCAGTACCTCTCCTCTCTCTCTCTCTCTCTCTCT 3675  
Db 361 CGACCTCTGTGGCGGCTGTACCTCAGTACCTCTCCTCTCTCTCTCTCTCTCTCTCT 420  
Qy 3676 GGGACAGGCCCAACCCAGGACAGGATTCCTTGAGGCCACAGAGGAGCACCAGAGGAA 3735  
Db 421 GGGACAGGCCCAACCCAGGACAGGATTCCTTGAGGCCACAGAGGAGCACCAGAGGAA 480  
Qy 3736 GATCTGTAAGTAGGCCCTTTGTTAGAGTCTCCAAAGGTTTCTCAGTCTGAGGCTCTCA 3795  
Db 481 GATCTGTAAGTAGGCCCTTTGTTAGAGTCTCCAAAGGTTTCTCAGTCTGAGGCTCTCA 540  
Qy 3796 CACACTCCCTCTCTCTCCCGGCTGTGGGTCTTCAATTTGCCAGCTCTCTGCCACACTCT 3855  
Db 541 CACACTCCCTCTCTCTCCCGGCTGTGGGTCTTCAATTTGCCAGCTCTCTGCCACACTCT 600  
Qy 3856 GCCTGTCTGCTGACGAGTGCATCATGTCTCTTGACAGAGGAGTCTGCACTCAAGGCC 3915  
Db 601 GCCTGTCTGCTGACGAGTGCATCATGTCTCTTGACAGAGGAGTCTGCACTCAAGGCC 660  
Qy 3916 TGAGGAAGCCTTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGAGGCTGCCAC 3975  
Db 1 TGAGGAAGCCTTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGAGGCTGCCAC 3975

[illegible]

RESULT	8
Q85435	
ID	Q85435 standard; DNA; 2420 BP.
AC	Q85435;
DT	09-OCT-1995 (first entry)
DE	Human melanoma antigen MAGE-1.
KW	Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
KW	HLA-restricted cytotoxic T-lymphocyte activity; ss.

OS	Homo sapiens.	:	Location/Qualifiers
PH	key	:	626. .1555
FT	cds	:	/*tag= a

FI PN WO9504542-A

PD 16-FEB-1995

PD 10-FEB-1993  
.PF 02-AUG-1994

02-AUG-1994  
06-AUG-1993PR 06-AUG-1993  
PA (CYTE-) CYTEPA (CYTE-) CYT  
PT Fikes JD.

PI Flkes JD,  
DR WPT: 95-0908

DR: \_\_\_\_\_  
 WP: \_\_\_\_\_  
 P-PSDB: \_\_\_\_\_

p-PSDB; R709

PT Human meland

PT stimulating

Example 1; I

Q85435 encoded

CC to produce t

CC R70969. The...

CC engender a

CC MAGE-1 antigens. Compsus. containing these peptides can be  
CC administered, as a vaccine to patients susceptible to MAGE  
CC associated tumours, e.g. melanomas.  
SQ Sequence 2420 BP; 562 A; 582 C; 677 G; 599 T;

Query Match 42.4%; Score 2408; DB 1; Length 2420;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2419; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy 3256 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGACAGAGGGGTCATCC 3315
Dy 1 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGACAGAGGGGTCATCC 60
Qy 3316 ACTGCATCAGAGTGGGGATGTACAGAGTCCAGGCCACCCCTCCTGGTAGCACTGAGAAGC 3375
Dy 61 ACTGCATCAGAGTGGGGATGTACAGAGTCCAGGCCACCCCTCCTGGTAGCACTGAGAAGC 120
Qy 3376 CAGGCTGTGTTGGGTCCTGCACCCCTGAGGCCCGGTGGATTCCTTCCTGGAGCTCCA 3435
Dy 121 CAGGCTGTGTTGGGTCCTGCACCCCTGAGGCCCGGTGGATTCCTTCCTGGAGCTCCA 180
Qy 3436 GGAACCCAGGAGTGGGCTTGGTCTGAGACAGTATCCTCAGGTCACAGACAGAGGATG 3495
Dy 181 GGAACCCAGGAGTGGGCTTGGTCTGAGACAGTATCCTCAGGTCACAGACAGAGGATG 240
Qy 3496 CACAGGGTGTGCCAGCAGTGAATGTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCA 3555
Dy 241 CACAGGGTGTGCCAGCAGTGAATGTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCA 300
Qy 3556 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACTCCCTACTGTCACTGTCTAGTAAAT 3615
Dy 301 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACTCCCTACTGTCACTGTCTAGTAAAT 360
Qy 3616 CGACCTCTGCTGGCGGCTGTACCTGTAGTACCTCTCAGTTCCTCTCAGGTTTCAG 3675
Dy 361 CGACCTCTGCTGGCGGCTGTACCTGTAGTACCTCTCAGTTCCTCTCAGGTTTCAG 420
Qy 3676 GGGACAGGCCAACCCAGAGGACAGGATTCCTGGAGGCCACAGAGGACACCAAGGAGAA 3735
Dy 421 GGGACAGGCCAACCCAGAGGACAGGATTCCTGGAGGCCACAGAGGACACCAAGGAGAA 480
Qy 3736 GATCTGTAAGTAGGCTTTGTAGAGTCTCAAGGTTCCAGTTCAGTCTGAGGCTCTCA 3795
Dy 481 GATCTGTAAGTAGGCTTTGTAGAGTCTCAAGGTTCCAGTTCAGTCTGAGGCTCTCA 540
Qy 3796 CACACTCCCTCTCTCCAGGCTGTGGTCTTCAATGCCAGCTCCTGCCACACTCCT 3855
Dy 541 CACACTCCCTCTCTCCAGGCTGTGGTCTTCAATGCCAGCTCCTGCCACACTCCT 600
Qy 3856 GCCTGCTGCCCTGACGAGTCACTATGTCCTTGAGCAGAGGAGTCTGCACCTGCAAGCC 3915
Dy 601 GCCTGCTGCCCTGACGAGTCACTATGTCCTTGAGCAGAGGAGTCTGCACCTGCAAGCC 660
Qy 3916 TGAGGAAGCCCTTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCAC 3975
Dy 661 TGAGGAAGCCCTTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCAC 720
Qy 3976 CTCCTCTCTCTCTCTCTGTCCTGGGACCCCTGGAGAGCTGCCACTGCTGGGTCAAC 4035
Dy 721 CTCCTCTCTCTCTCTCTGTCCTGGGACCCCTGGAGAGGCTGCCACTGCTGGGTCAAC 780
Qy 4036 AGATCTCTCCAGAGTCTCTCAGGAGGCTCCGCGCTTTCCCACTACCATCACTCACTCG 4095
Dy 781 AGATCTCTCCAGAGTCTCTCAGGAGGCTCCGCGCTTTCCCACTACCATCACTCG 840
Qy 4096 ACAGAGCAACCCAGTAGGCTTCCAGAGCCGTCAGAGAGGGGCCAACACCTCTTG 4155
Dy 841 ACAGAGCAACCCAGTAGGCTTCCAGAGCCGTCAGAGAGGGGCCAACACCTCTTG 900
Qy 4156 TATCTGGAGTCTCTGTTCCAGCAGTAATCACTAAGAGCTGGCTGATTTGGTGT 4215
Dy 901 TATCTGGAGTCTCTGTTCCAGCAGTAATCACTAAGAGCTGGCTGATTTGGTGT 960
```

```
Qy 4216 TCTCTCTCTCAAAATATCAGCCAGGAGCCAGTCACAAAGCAGAAATCTGTGAGAGTGT 4275
Dy 961 TCTCTCTCTCAAAATATCAGCCAGGAGCCAGTCACAAAGCAGAAATCTGTGAGAGTGT 1020
Qy 4276 CATCAAAAATTTACAAGCACTGTTTCTCTGAGATCTTGGGCAAGCCCTCTGAGTCTTGCA 4335
Dy 1021 CATCAAAAATTTACAAGCACTGTTTCTCTGAGATCTTGGCAAGCCCTCTGAGTCTTGCA 1080
Qy 4336 GCTGCTCTTTGGCATTCACGTGAAGAGCAGACCCACCCGCCCTCTCTGTCCTTGT 4395
Dy 1081 GCTGCTCTTTGGCATTCACGTGAAGAGCAGACCCACCCGCCCTCTCTGTCCTTGT 1140
Qy 4396 CACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4455
Dy 1141 CACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
Qy 4456 AGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4515
Dy 1201 AGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
Qy 4516 GGAATCTGGGAGGAGCTGATGATGGAGTCTATGATGGAGGAGGAGCAGTGCCTA 4575
Dy 1261 GGAATCTGGGAGGAGCTGATGATGGAGTCTATGATGGAGGAGGAGCAGTGCCTA 1320
Qy 4576 TGGGAGGCCAGGAGCTGCTCACCCAAAGATTGGTGAGGAGGAGGAGGAGGAGGAGGAGG 4634
Dy 1321 TGGGAGGCCAGGAGCTGCTCACCCAAAGATTGGTGAGGAGGAGGAGGAGGAGGAGGAGG 1380
Qy 4635 GCAGGTGCCGAGCAGTGTATCCCGCACGCTATGATGCTGTGGGTCCTCAAGGGCCCTGCG 4694
Dy 1381 GCAGGTGCCGAGCAGTGTATCCCGCACGCTATGATGCTGTGGGTCCTCAAGGGCCCTGCG 1440
Qy 4695 TGAACACAGCTATGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4754
Dy 1441 TGAACACAGCTATGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Qy 4755 TTTCTTCCCATCCCTGGCTGAAGCAGCTTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4814
Dy 1501 TTTCTTCCCATCCCTGGCTGAAGCAGCTTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
Qy 4815 AGTTGACGCAAGGCCAGTGGAGGGGAGTGGGCCAGTGCACCTTCCAGGGCCGCTGCC 4874
Dy 1561 AGTTGACGCAAGGCCAGTGGAGGGGAGTGGGCCAGTGCACCTTCCAGGGCCGCTGCC 1620
Qy 4875 AGCAGCTTCCCTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4934
Dy 1621 AGCAGCTTCCCTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
Qy 4935 GTGTTCTCAGTGTAGTGTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCT 4994
Dy 1681 GTGTTCTCAGTGTAGTGTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCT 1740
Qy 4995 TTTGGAATTTTCAAAATGTTTTTTTTTAAAGGATGTTGAAATGAACCTTCAGCATCCCAAGT 5054
Dy 1741 TTTGGAATTTTCAAAATGTTTTTTTTTAAAGGATGTTGAAATGAACCTTCAGCATCCCAAGT 1800
Qy 5055 TTATGAATGACAGCAGTACACACTTCTGTGTATATAGTTTAAAGGATGAGTCTGTGT 5114
Dy 1801 TTATGAATGACAGCAGTACACACTTCTGTGTATATAGTTTAAAGGATGAGTCTGTGT 1860
Qy 5115 TTTATTTCAGATTTGGAAATCCATTTCTATTGTAATTTGGATTAATACACAGCTGGAAT 5174
Dy 1861 TTTATTTCAGATTTGGAAATCCATTTCTATTGTAATTTGGATTAATACACAGCTGGAAT 1920
Qy 5175 AAGTACTTAGAAAATGTGAAAATGAGCAGTAAAAATAGATGAGATAAGAACTAAAGAAAT 5234
Dy 1921 AAGTACTTAGAAAATGTGAAAATGAGCAGTAAAAATAGATGAGATAAGAACTAAAGAAAT 1980
Qy 5235 TAAGATAGTCAATCTCTGCTTATACCTCAGTCTATCTGTAAAAATTTTTTAAGATAT 5294
Dy 1981 TAAGATAGTCAATCTCTGCTTATACCTCAGTCTATCTGTAAAAATTTTTTAAGATAT 2040
```

QY 5295 ATGCATACCTGGATTCTTCTGGCTTCTTTGAGATGTAAGAGAAATTAATCTGAATAAA 5354  
 |||||  
 Db 2041 ATGCATACCTGGATTCTTCTGGCTTCTTTGAGATGTAAGAGAAATTAATCTGAATAAA 2100  
 |||||  
 QY 5355 GAATTCCTCTCTTCTCACTGGCTCTTTCTCTCCATGCACTGAGCATCTGCTTTTGGAA 5414  
 |||||  
 Db 2101 GAATTCCTCTCTTCTCACTGGCTCTTTCTCTCCATGCACTGAGCATCTGCTTTTGGAA 2160  
 |||||  
 QY 5415 GGCCTGGSTTAGTAGTGGAGATGCTAAGTAAAGCCAGACTCATACCCACCATAGGTC 5474  
 |||||  
 Db 2161 GGCCTGGSTTAGTAGTGGAGATGCTAAGTAAAGCCAGACTCATACCCACCATAGGTC 2220  
 |||||  
 QY 5475 GTAGAGCTTAGAGCTGAGTACGTAATCGAGGTGGCAAGATGCTCTTAAGATGTAG 5534  
 |||||  
 Db 2221 GTAGAGCTTAGAGCTGAGTACGTAATCGAGGTGGCAAGATGCTCTTAAGATGTAG 2280  
 |||||  
 QY 5535 GGAAGAGTGAGAGAGGGGTGAGGGTTCGGGCTCGGGTGAGAGTGGTGAAGTCAATG 5594  
 |||||  
 Db 2281 GGAAGAGTGAGAGAGGGGTGAGGGTTCGGGCTCGGGTGAGAGTGGTGAAGTCAATG 2340  
 |||||  
 QY 5595 CCCTGAGCTGGGGCATTTTGGGCTTTGGGAACTGCAGTTCCTTCTGGGGGAGCTGATTG 5654  
 |||||  
 Db 2341 CCCTGAGCTGGGGCATTTTGGGCTTTGGGAACTGCAGTTCCTTCTGGGGGAGCTGATTG 2400  
 |||||  
 QY 5655 TAATGATCTGGGTGATCC 5674  
 |||||  
 Db 2401 TAATGATCTGGGTGATCC 2420

## RESULT 9

Q72478  
 ID Q72478 standard; DNA: 4157 BP.  
 AC Q72478;  
 DT 22-JUN-1995 (first entry)  
 DE Tumour rejection antigen MAGE-2 gene.  
 KW Tumour rejection antigen; melanoma antigen-2; MAGE-2; MAGE-3;  
 KW cancer; cytolytic T cells; antigen D; human leucocyte antigen;  
 OS Homo sapiens.  
 FH Key  
 FT cds  
 FT /\*tag= a  
 PN W09423031-A.  
 PD 13-OCT-1994.  
 PF 17-MAR-1994; U02877.  
 PR 26-MAR-1993; US-037230.  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PI Boon-faller T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;  
 DR WPI: 94-333192/41.  
 PT New tumour rejection antigen precursor MAGE3 - useful in  
 treatment and diagnosis of cancer  
 PS Example 32; Page 61; 105pp; English.  
 CC Q72478 is the gene which contains the coding sequence for melanoma  
 antigen-2 (MAGE-3). Another melanoma antigen precursor. Melanomas  
 characterised by the expression of MAGE-3 can be detected, or  
 monitored, by contacting a test sample with an agent that can  
 recognise MAGE-3. The melanoma can be treated by the administration  
 of cytolytic T cells specific for the complex of antigen D (the  
 mature rejection antigen derived from MAGE-3) and a human leucocyte  
 antigen (esp. HLA-A1).  
 SQ Sequence 4157 BP; 953 A; 1134 C; 1185 G; 885 T;

Query Match 36.3%; Score 2058.8; DB 1; Length 4157;  
 Best Local Similarity 75.3%; Pred. No. 0;  
 Matches 3172; Conservative 0; Mismatches 837; Indels 201; Gaps 41;

QY 1351 CCCATCGCCTCCCCCATCTTGGCAGAACTCCGGT-TGCCCTCTCTCAACCCAGGGAAG 1409  
 |||||  
 Db 1 CCCATCCAGATCCCATCCGGCAGAAATCCGGTTCACCTCTGCCGTGAACCCAGGGAAG 60  
 |||||  
 QY 1410 CCCTGGTAGGCCCGATGTGAAACCACTGACTGTAACCTCACAGATCTGAGAAAGCCAGG 1469  
 |||||

Db 61 TCACGG--GCCCGATGTGACGCCACTGACTTGACATTTGGAGGTGAGAGGACAGGCGA 118  
 |||||  
 QY 1470 TTCAATTAATGTTCTTCTGAGGGCGCTTGAGATCCACTCAGGGGAGTGTGTTTAGGCTCT 1529  
 |||||  
 Db 119 TTC-----TCGCCCTGAGCAACGGCTGACGTGCGGAGGGAAGCAGGCGAGGCTCC 172  
 |||||  
 QY 1530 GTGAGAGGCAAGGTGAGATGCTGAGGGAGGACTGAGGAGGCACACACCCAGGTAGATG 1589  
 |||||  
 Db 173 GTGAGAGGCAAGGTGAGATGCTGAGGGAGGACTGAGGCGGCTCACCACAGACAGAGG 232  
 |||||  
 QY 1590 GCCCAAAATGATCAGTACACCCCTGCTGCCAGCCCTGGACCACCCCGCCAGGACAGA 1649  
 |||||  
 Db 233 GCCCCCAATTAATGAGCGCTGCTGCTGCCGGCTGGACACCTCGCAGGGGAAGA 292  
 |||||  
 QY 1650 TGTCTCAG:-----CTGACACACCCCGCTCCCGTCCCACTGCCACTTAAACCCACAGG 1702  
 |||||  
 Db 293 CTTCTCAGGCTCAGTCGCCACCACTCACCCCGCCACCCCGCGCTTAAACCGCAGGG 352  
 |||||  
 QY 1703 CAATCTGTAGTCAATAGCTTA-TGTGACCGGGGACAGGTTGTCAGGAGAGGAGGCCCA 1761  
 |||||  
 Db 353 AACTCTGGCGTAAGAGCTTGTGTACCAAGGCGAGGCTGGTTAGAAAGTGC----- 404  
 |||||  
 QY 1762 GGCATCAAGGTCCAGATCCCGCCCGGCAATAGGTCAGGACCCCTGGGAGGGAACCTGAGG 1821  
 |||||  
 Db 404 -----TCAGGGCCAGACTCAGCCAGGAATCAAGGTCAAGACCCCAAGAGGGGACTGAGG 459  
 |||||  
 QY 1822 TTCCCCACCCACACCTGCTCTCATCTCC----- 1852  
 |||||  
 Db 460 CAACCCACCCCTACCTACCTACCAATCCCATCCCAACACCAACCCACCCCATCC 519  
 |||||  
 QY 1852 -----ACGCCACCCCACTCACATTCCTACCTACCCCTACCCCAACCTCATCTGT 1906  
 |||||  
 Db 520 CTCAAAACCAACCCCAACCCCACTTCCCATCTCTCTCCCAACCACTCATCTGG 579  
 |||||  
 QY 1907 CAGAAAT-----CCCTGCTGTCAACCCAGGAGCCAGGGAATGGCGGCCAGGCA 1956  
 |||||  
 Db 580 CAGAAATCCGGCTTTGCCCTGCAATCAACCCAGGAAGTCCGGGAATGGCGGCCAAGCA 639  
 |||||  
 QY 1957 CTCGGATCTTGACGTCCCACTCAGGGTCTGATGAGGGAAGGG-----CTTGAAC 2008  
 |||||  
 Db 640 CGCGGATCTTGACGTCTACATGTACGG-CTAAGGAGGGAAGGGTTGGGTCTCTGTAGT 698  
 |||||  
 QY 2009 AGGGCTCAGGGAGCAGAGGAGGCG-----CCTACTCGAGATGAGGAGGCGCTCAGA 2063  
 |||||  
 Db 699 ATGGCTTTGGGATGAGAGGAAGGCGCCAGGCTCTCTGGAAGACAGTGGAGTCTTAGG 758  
 |||||  
 QY 2064 GGACCCAGCACCCTAGGA-----CACCGCACCCTGTCTGAGACTGAGGCTGCCAC 2114  
 |||||  
 Db 759 GGACCCAGCATGCCAGGACAGGGGCGCCACTGTACCCCTGTCTCAAACTGAGCCACCTT 818  
 |||||  
 QY 2115 TTCTGSCCTCAAGAAATCAGAAAGATGGGACTCAGATTCATGGGGGTGGGACCCAGGCC 2174  
 |||||  
 Db 819 TCATTCAGCCGAGGGAATCCTAGGATGAGACCCACTTCAGGGGTTGGGCGCCAGCCT 878  
 |||||  
 QY 2175 TCAAGGCTTACCGGAGGAAGAGGAGGAGGACTCAGGGGACCTTTGGAATCCAGATCAG 2234  
 |||||  
 Db 879 GCGAGGAGTCAAGGGGAGGAAGAGGAGGAGGACTGAGGGGACCTTTGGAGTCCAGATCAG 938  
 |||||  
 QY 2235 TGTGAGACTCGGCCCTGAGAGGTCCAGGCGAGGTGGGCCACATATATGGCCCATATTTCTTG 2294  
 |||||  
 Db 939 TGGCAACCTTGG-GCTGGGGGATCTTGGGCGAGTGGCGGAATGTGCCCGCTGCTCATTTG 997  
 |||||  
 QY 2295 CATCTTTGAGGTGAC----AGCACAGAGCTGTGCTGAGAAGTGGGGCTCAGGTCAC 2350  
 |||||  
 Db 998 CACCTTCAGGGTGACAGAGAGTTGAGGGCTGTGGTCTGAGGGCTGGGACTTCAGGTGAGC 1057  
 |||||  
 QY 2351 AGAGGAGGAGTTCAGAGATCCATATGGCCCAAGATGTGCCCTTCATGAGGAGTGGG 2410  
 |||||  
 Db 1058 AGAGGAGGAATCCAGGATCTGCCGACCCCAAGGTGTGCCCTTCATGAGGACTCCCC 1117  
 |||||  
 QY 2411 ATATCCCCGGCTCAGAAAGAGGAGCTCCACACAGTCTGGCTGTCCCTTTTAGTAGCTC 2470  
 |||||



Db	1118	ATACCCCGCCGACGAAAGGGATGCCACAGACTCGGAAGTAAATTTGTTCTAGCTC	1177
Qy	2471	TAGGGGACACATCAGGATGCGGTATGTTCATTCTCACTTGTACACAGCAGGAA	2530
Db	1178	TGGGAAACCTGATCAGGATGSCCTTAAGTACAATCTCATTTGTACCACAGCAGAG	1237
Qy	2531	GTTCGGGGCCCTCAGGAGATGGGGTCTTGGGTAAAGGGGAGTCTACTCATGTCA	2590
Db	1238	GTTCGGGAAACCTCAGGAGATAAGGTGTGGTGTAAAGAGGAGCTGTCTCATTTCA	1297
Qy	2591	GGGAATTGGGGTTGAGGAAGCACAGGCCCTGGCAGGAATAAAGTANGTCAGACACA	2650
Db	1298	GGGGTTTCCCTTGAGAAAGGCGACTCCCTGGCAGGAGTAAAGTAGTAACCCAGG	1357
Qy	2651	AGCTATTGGAATCCACACCCACAGAACAAAGGGGTACGCCCTGGACACTCACCCAGGA	2710
Db	1358	AGCCATCATACGTTACCTTAGAACCAAGGGGTACGCCCTGGACACGCCAGTGGG	1417
Qy	2711	T-----GTGGCTCTTTTTTCACCTCTGTTCACAGATCTGGGCGAGGTGAGGACCTCAT	2763
Db	1418	TAACAGGATGTGSCCCTCTCTCACTGTCTTTTCCAGATCTCAGGGAGTTGATGACCTTGT	1477
Qy	2764	TCTCAGAGGATGACTCAGGTCAAGCTAGGGACCCCCATCTGGTCTAAGACACAGACGGTC	2823
Db	1478	TTTCAAGAGGTGACTCAGTCAACACAGGG--CCCTCTGGTCGACAGATGCAAGTGGTT	1534
Qy	2824	CCAGGATCTGCCATCGTTTCGGGTGAGGAACATGAGGAGGACGTAGGGGTACCCAGGAC	2883
Db	1535	CTAGGATCTGCCAAGCATCCAGGTGGAGGCTGAGGTAGGATTGAGGGTACCCCTGGGC	1594
Qy	2884	CAGAACACTG-AGGGAGACTGCACAGAATCAGCCCTGCCCTCTGTCAACCCAGAGAG	2942
Db	1595	CAGAATGCAGAGGGGGCCCATAGAAATCTGCCCTGCCCTCGGTTACTTCAGAGAC	1654
Qy	2943	CATGGGCTGGCGCTGCCGAGTCTTCGGTATCTCTGGGATCATGTGTCAGGAC	3002
Db	1655	CCTGGGACGGGTGTCAAGTGAAGTCCCTCCATTAT-CTGGGATCTTTGATGTCAAGGAA	1713
Qy	3003	GGGAGGCTTGTGTGAGAAGGCTCGCTCAGGTCAGTCAGAGGAGCGTCCAGAGCCCT	3062
Db	1714	GGGAGGCTTGGTCTGAAGGGCTTGAAGTCAGTCAAGTGAAGGAGGGTCTCAGGCCCT	1773
Qy	3063	GCCAGGATCAAGGTGAGGACCAAGCGGCACCTCACCCAGGACACATTAAATTCATGA	3122
Db	1774	GCCAGGATGGAGCTGAGGACCAAGCGGACTCGTCACCCAGGACACCTGGACTCCAATGA	1833
Qy	3123	ATTTGATATCTTCTCCCTTCCCAGGACCTAGGACAGTGTGGCCAGATGTTGT	3182
Db	1834	A-TTTGACATCTCTGTCTCGTCGGGAGGACCTGTGTACGTATGGCAGCATGTGGT	1892
Qy	3183	CCCTC-CTGTCTTCCATTCTTATCATGATGTGAACCTCTG-ATTGGATTCTCA	3239
Db	1893	CCCTCTATCTCTGTACCATATCAGGATGTAGTCTTGACATGAGAGATTCTCA	1952
Qy	3240	GACCAGCAAAAGGCGGATTCAGGCCCTTGGCAGGAAATATTAAGGGCCCTCGCTGAGA	3299
Db	1953	AGCCACAAAGGTTGGAT-TAGGCCCTACAAAGGAGAAAGTTCAGGGCCCTGAGTGAGC	2011
Qy	3300	ACAGAGGGGTATTCATGCTGATAGAGTGGGATGTCAAGATGTCAAGCCACCTTCCT	3359
Db	2012	ACAGAGGGACCTTCCACCAAGTAGTAGTGGGGACCTCAGCGAGTCTGGCCAAACCTGCT	2071
Qy	3360	GGTAGCACTGAGAACCGGCTGTGCTTGGGTTCTGCACCTCAGGGCCCGTGGATTCC	3419
Db	2072	GAGACTTCTGGGAATCCGTGGCTGTGCTTTGCACTGTGCACACTGAAGGCCGTGCATTC	2131
Qy	3420	TCTTCT-----GGAGTCTCAGGAACCGAGCAGTAGGCGCTTGGTCTGAGACAGTATC	3472
Db	2132	TCTCCAGGAATCAGGAGCTCCAGAACCGAGCAGTAGGCGCTTGGTCTGAGTCACT-GC	2190
Qy	3473	CTCAGTCTACAGACAGGAGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATG	3532
Db	2191	CTCAGTCTACAGACAGGGGACGACAGTGCACACTGAAGTCTTGGTGGTGGTGAATG	2250



Qy	4584	CCAGGAAGCTGCTCACCACCAAGATTGGTGCAGGAAAAAGTACTCGAGTA-CGGCAGGTCG	4584
Db	3322	CCAGGAAGCTGCTCATGCAAGATCTGGTGCAGGAAAAAGTACTCGAGTACCGGCAGGTCG	3381
Qy	4643	CGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGTCCCAAGGGCCCTCGCTGAACCA	4702
Db	3382	CCGGCAGTGATCTGCTGCTACGAGTTCCTGTGGGTCCCAAGGGCCCTCATTTGAACCA	3441
Qy	4703	GCTATGTGAAAGTCCCTTGAGTATGTGATCAAGGTCAGTGCAGAGAGTTCGGTTCCTTC	4762
Db	3442	GCTATGTGAAAGTCCCTGCACATACACTAAAGATCGGTGGAGACCTCACATTCCTTACC	3501
Qy	4763	CATCCCTCGGTGAAGCAGCTTTGAGAGAGGAGAGAGGGAGTCTGAGCATGAGTTGCAG	4822
Db	3502	CACCCCTGCATGAACGGCTTTGAGAGAGGGAGAGAGTGTAGTCTCAGCACATGTTGCAG	3561
Qy	4823	CCAAGGCCAGTGGGAGGGGAGCTGGCCACAGTCCACCTTCCAGGGCCGGTCCAGCAGCTT	4882
Db	3562	CCAGGGCCAGTGGGAGGGGCTGTGGCCACAGTGCACCTTCCAGGGCCCCATCCCATTAGCTT	3621
Qy	4883	CCCTCGCTCGTGTGACATGAGGCCATTCTTT--CACCTCTGAAGAGAGCGGTGAGTCTTC	4940
Db	3622	CCACTGCTCGTGTGATATGAGGCCATTCTTCCTCTTTGAAGAGAGCAGTCAGCATCTT	3681
Qy	4941	TCAGTAGTAGGTTTCTGTCTATTTGGGTGACCTGGAGATTATCTTTGCTCTCTTTTGGGA	5000
Db	3682	TTAGCAGTGAGTTTCTGTTCTGTGTGATGACCTTTGAGATTATCTTCTTTCCGTGTGGA	3741
Qy	5001	ATTGTTCAAAATGTTTTTTTTTAAAGGATGGTTGGAATGAACCTTCAGCATCCAAAGTTTATGA	5060
Db	3742	ATTGTTCAAAATG-TTCCCTTTTAAACAAATGGTTGGATGAACCTTCAGCATCCAAAGTTTATGA	3800
Qy	5061	ATCAGACGAGTCACAC--AGTTCGTGTATATAGTTTTAAAGGGTAAGAGTCTTGTGTTTTA	5118
Db	3801	ATCAGCAGTAGTCACACATAGTCTCTTATATAGTTTAAAGGGTAAGAGTCTGTTTTTTA	3860
Qy	5119	TTCAGATTGGGAAATCCATTCTATTTTGTGAAATG--GGATAATAACAGCAGTGGAAATA	5176
Db	3861	TTCAGATTGGGAAATCCATTCCATTTGTGAGTCTGCACATANTACAGCAGTGGAAATAT	3920
Qy	5177	GTACTTAGAAATGT----GAAAAATGACGAGTAAATAGATGAGATAAAGAACTTAAAGAA	5232
Db	3921	GTATTTGCCCTATATCTGTGAACGAATAGCAGTAAATACATGATACAAGGAAC-----	3974
Qy	5233	ATTAAGAGATAGTCAATCTTCGCTTATACCTCAGTCTATTCTGTAAATTTTAAAGAT	5292
Db	3974	-TCAAAGATAGTTAAATCTTCGCTTATACCTCAGTCTATTATGTAAAA---TTAAAAAT	4029
Qy	5293	ATATGATACCTGGATTTCTTGGCTCTCTTGAGAACTGAAGAAATTAATCTCGAATA	5352
Db	4030	ATGTGTATGTT-----TTTGCTCTCTTGAGAAATGCAAAAGAAATTAATCTCGAATA	4080
Qy	5353	AAGAAATCTTCCTGTCTACTGGCTCTTTTCTCCATGCACTGAGCATCTGCTTTTTGG	5412
Db	4081	---AATCTTCCTGTCTACTGGCTCATTTCTTTACCATTCACTCAGCATCTGCTCTGTGG	4137
Qy	5413	AAGCCCTCGG	5422
Db	4138	AAGGCCCTGG	4147

## RESULT 10

RESULTS	10
Q32353	
ID	Q32353 standard; DNA; 4157 BP.
AC	Q32353.
DT	22-APR-1993 (first entry)
DE	MAGE-2 gene.
KW	melanoma antigen; MAGE TRA; melanoma antigen tumor rejection antigen;
KW	tumor rejection antigen precursor; MAGE; antigen E; gene family; ss.
OS	Homo sapiens.
FH	Key
FT	Location/Qualifiers
FT	2958..3542
FT	cds
FT	/*tag= a



Db	3801	ATGACAGT	AGTACACACATAGTCTCTTTATATATAGTTT	AGGGGT	AAGAGTCTCTGTTTTTTA	3866
QY	5119	TTCAGAT	TTGGGAATCCATTTCTATTTCTGTAATTC	-	GGATAATAACAGCAGTGGAAATAA	5176
Db	3861	TTCAGAT	TTGGGAATCCATTTCTATTTCTGTAATTC	TTT	GGAGTGTACATTAATACAGCAGTGGAAATAT	3920
QY	5177	GTACT	TAGAAATGT	----	GAAATAGCAGTAAATAGATGAGATAAAGAACTAAAGAA	5232
Db	3921	GTATTT	GGCTATATTTGTGAACGAATTAGCAGTAAATACATGATACAGGAAC	-----	3974	
QY	5233	ATTAAGAC	ATAGTCAATCTTCGCTTTATACCTTCAGTCTATTCTCTAATAATTTTAAAGAT		5292	
Db	3974	-TCAA	AAGATAGTAAATTTCTGCTTTATACCTTCAGTCTATTCTATGTAAAA	---	TTAAAAAT	4029
QY	5293	ATATGAT	ACCTGGATTTCCCTTGGCTTTCTTTGAGAATCTAAGAGAAATTTAAATCTGAATA		5352	
Db	4030	ATGTG	TATGTT-----	TTTGCTTCTTTGAGAATSCAAAGAAATTTAAATCTGAATA	4080	
QY	5353	AGAAT	TTCTCTGTTCACTGGCTCTTTTCTTCCTCAGTCTGAGCATCTGCTTTTGG		5412	
Db	4081	----	AATCTCTGTTCACTGGCTCTATTCTTTACCATCTACTCAGCATCTGCTCTGTGG		4137	
QY	5413	AAGGCC	CTCGS	5422		
Db	4138	AAGGCC	CTCGG	4147		
RESULT 11						
X26974						
ID	X26974	standard; cDNA; 4204 BP.				
AC	X26974;					
DT	25-JUN-1999	(first entry)				
DE	CNA encoding MAGE-3 polypeptide.					
KW	MAGE-3 tumour associated gene; human leucocyte antigen Class II;					
KW	autologous CD4+ cell; MAGE-3 related disease; cancer; melanoma;					
KW	osteosarcoma; leukemia; carcinoma; ss.					
OS	Homo sapiens.					
FT	Key	Location/Qualifiers				
FT	CDS	2465..3409				
FT		/*tag= a				
FT		/product= "MAGE-3"				
PN	W09914326-Al.					
PD	25-MAR-1999.					
PR	04-SEP-1998; U18601.					
PR	12-SEP-1997; US-928615.					
PA	(LUDW-) LUDWIG INST CANCER RES.					
PA	(UYVR-) UNIV VRIJE BRUSSEL					
PI	Boon-Faller T; Chaux P, Corthals J, Heirman C,					
PI	Luiten R, Stroobant V, Thielemans K, Van Der Bruggen P;					
DR	WPI; 99-244031/20.					
DR	P-PSDR; Y01720.					
PT	Isolated peptides that bind to human leucocyte antigen class II					
PT	molecules					
PS	Example 5; Page 65-67; 88pp; English.					
CC	The present sequence represents the MAGE-3 tumour associated gene.					
CC	Peptides that bind human leucocyte antigen (HLA) Class II molecules					
CC	can be derived from the MAGE-3 protein. These peptides and					
CC	autologous CD4+ cells that bind to a complex of MAGE-3 peptide					
CC	and HLA Class II, are used to treat MAGE-3 related diseases,					
CC	particularly cancers (e.g. melanoma, osteosarcoma, leukemia and					
CC	various forms of carcinoma). The peptides are also used to produce					
CC	specific antibodies. Detection of the peptides, e.g. in binding					
CC	assays, particularly with antibodies, is used for diagnosis of such					
CC	diseases.					
QY	Sequence	4204 BP;	944 A;	1144 C;	1223 G;	893 T;
Query Match						
Best Local Similarity		31.1%;		Score 1762.6;		DB 1; Length 4204;
Matches 2315;		Conservative		0;		Mism









[illegible]

D	b	1139	CCACAGCAACACTACACCCCTTGTACCGCTGGGCGCTTTCCTATGATGGCCGTGCTGG	1198
Q	y	4431	GTGATATCAGATCATGCCCAAGACAGCGCTCCCTGATAAATTGCTCCTGGTCATGATTCGAA	4490
D	b	1199	GTAATAATCAGATCTTCCCAAGACAGCGCTTCTGATTAATCGTCTGGGCACAATTCGAA	1258
Q	y	4491	TGGAGGCGGCCCATGCTCTCGTAGGAGGAAATCTGGAGAGAGCTGAGTGTGATGGAGGTGT	4550
b	b	1259	TGGAGGCGACACGCGCTCTGAGGAGGAAATCTGGAGAGAGCTGGGTGCTGATGGGGGTGT	1318
Q	y	4551	ATGATGGGAGGGAGCACATGCTCTATGGGAGGCCCAGGAAGCTGCTCACCCACAGATTGG	4610
D	b	1319	ATGATGGGAGGGAGCACATGCTCTATGGGAGGCCCAGGAAGCTGCTCACCCACAGATTGGG	1378
Q	y	4611	TGCAGGAAAAAGTACCTGGAGTA-CGCGAGGTGCGGACAGATGATCCGCGAGCTATCAGT	4669
D	b	1379	TGCAGGAAAACTACCTGGAGTACCGGCAAGTACCGCGAGTAATCTCGCGGCTATCAGT	1438
Q	y	4670	TCCTGTGGGTCCTCAAGGGCCCTCGCTGAAACAGCATATGTGAAAGTCTCTTCAGTATGTGA	4729
D	b	1439	TCCTGTGGGTCCTCAAGGGCTCGCTGAAACAGCATATGTGAAAGTCTCTTCAGTATGTGA	1498
Q	y	4730	TCAAGGTCACTGCAAGAGTTCGCTTTTCTTCCCATCCCTCGGTGAACAGCTTTTGAGAG	4789
D	b	1499	TCAGGGTCAATGCAAGAGTTCGCAATTCACCATCCCTCGCTGAAAGAGCTTTGTTAG	1558
Q	y	4790	AGGAGGAGGAGGAGTCTGAGCATGATGCTGAGCCAAAGGCCAGTGGGAGG- ---GGACT	4845
D	b	1559	AGGAGGAGGAGGAGTCTGAGCATGATGCTGAGCCAGGCTTCCCTCGCTGTGACATCAGG	1618
Q	y	4846	GGGCGAGTGCACCTTCAGGGCGCGGTCCAGCAGCTTCCCTCGCTGTGACATCAGG	4905
D	b	1619	GGGCGAGTGCATCTACA- -GCCCTGTGCAGCAGCTTCCCTTGCTCGTGAACATCAGG	1676
Q	y	4906	CCCATTCTTCACTC- - -TGAAGAGAGCGGTGAGTGTCTTCAGTAGTAGGTTCTCTGTCT	4961
D	b	1677	CCCATTCTTCACTCCTTTGAAGAAAAAGTACAGTGTCTTAGTAGTGGGTTCTATTTT	1736
Q	y	4962	ATTGGGTGACTTGGAGATTATCTTTGTTCTCTTTTGAATTTGTTCAATGTTTTTTTTT	5021
D	b	1737	GTTGGATGACTTGGAGATTATCTCTGTTCCCTTTTACAATTTGTGAAATG-TTCCCTTTT	1795
Q	y	5022	AAGGATGGTGAATGAACCTTCAGCATCCAGATTTATGAATGACAGCAGTCCACACAGTTC	5081
D	b	1796	ATGGATGGTGAATTAACCTCAGCATCCAGTTTATGAATCGTAGTTAACGTATATGC	1855
Q	y	5082	TGTGTATATAGTTTAAGGTAAGAGTCTTGTTTATTCAGATTGGGAAATCCCATCTTA	5141
D	b	1856	TGTTAATATAGTTAGGATGAAGTCTTGTTTTTATTTCAGATTGGGAAATCCGTTCTTA	1915
Q	y	5142	TTTTGTAATTTGG- - -ATAATAACAGCAGTGGGAATAGTACTTAGAATGTGAAAAATG	5198
D	b	1916	TTTTGTAATTTGGGACATAATAACAGCAGTGGAGTAGTATTTAGAAGTGTG- - -AATT	1972
Q	y	5199	AGCAGTAAAAATAGATCAGATAAAGAACTAAAGAAATTAAGAGATAGTCAATTTCTGCGCT	5258
D	b	1973	CACCGTGAATAGGTGAGAT- -	2020
Q	y	5259	ATACCTCAGTCTATTCTGTAATAATTTTAAAGATATATGCATACCTGGATTTCCTTGCGCT	5318
D	b	2021	ATGCGCTCAGTCTATTCTGTAAAATTTAAAAATATATATGCATACCTGGATTTCCTTGCGCT	2080
Q	y	5319	TCTTTGAGATGAAGAGAAATTAATCTGAATAAAGAAATTTCTTCCTGTTCCACTGGCTCT	5378
D	b	2081	TC- - -GTGAAATGTAAGAGAAATTAATCTGTAATAAATAATTTCTTCTGTTAACTGGCTCA	2137
Q	y	5379	TTTCTTCTCATGCATGAGCATCTGCTTTTTTGGAGGCCCTGGGTAGTAGTGAGATG	5438
D	b	2138	TTTCTTCTCATGCATGAGCATCTGCTCTGTGGAGGCCCGCCAGGAATAGTAGTGAGATG	2197
Q	y	5439	CTAAGGTAAAGCCAGACTCATACCCACCATAGGGTCTGTAGAGTCTAGGAGCTGCAGTCCAC	5498









Search completed: December 11, 1999, 06:29:38  
Job time: 12965 sec

1999-12-11 06:29:38

**THIS PAGE BLANK (USPTO)**